DOCUMENT RESUME

ED 033 085 SP 003 213

The Emerging Research Worker and His Needs. Report of the Convocation on Educational Research. (7th, Nov. 13-15, 1966).

Educational Research Association of New York State.; New York State Univ. System, Albany.

Pub Date 66

Note-186p.

EDRS Price MF-$0.75 HC-$9.40

Descriptors-Disadvantaged Youth, Educational Administration, *Educational Research, *Educational Researchers, Instructional Media, *Professional Training, Programed Instruction, *Research Methodology, Staff Role, Student Characteristics, Teacher Education

This report contains 73 addresses and papers presented at a convocation on the place of educational research in the larger social setting, and, more specifically, on the role that the educational researcher can and will play. Major presentations focus on the topics of educational research and the accumulation of knowledge, curriculum modifications for the emerging research worker, needs of the researcher, and the question of whether the researcher shall be a computer scientist or a Renaissance man. Other talks consider such matters as research methodology, taxonomy, and the training of educational research workers. Papers presenting highlights of recent research deal with educational administration, Bayesian statistics, and curriculum and teaching. Symposia presentations include discussion of complex educational media, needed research in teacher education, and Title I of the Elementary and Secondary Education Act of 1965. Focusing on specific knowledge in the field, the 55 reports of current research (some of them abstracts) deal with such subjects as programed instruction, research related to disadvantaged children, concept formation, instructional mode, reading, teacher tenure laws, research methodology, and studies of high school and college student attitudes and personality characteristics. (JS)

ED0 33085

THE EMERGING RESEARCH WORKER AND HIS NEEDS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

REPORT OF THE SEVENTH CONVOCATION ON EDUCATIONAL RESEARCH

HELD NOVEMBER 13-15, 1966

SPONSORED BY
THE UNIVERSITY OF THE STATE OF NEW YORK
AND THE
EDUCATIONAL RESEARCH ASSOCIATION OF NEW YORK STATE

SP003213

The University of the State of New York
Regents of the University (with years when terms expire)

1968 Edgar W. Couper, A.F., LL.D., L.H.D., Chancellor..Binghamton
1970 Everett J. Penny, B.C.S., D.C.S., Vice Chancellor.White Plains
1978 Alexander J. Allan, Jr., LL.D., Litt.D............Troy
1973 Charles W. Millard, Jr., A.B., LL.D...............Buffalo
1972 Carl H. Pforzheimer, Jr., A.B., M.B.A., D.C.S.....Purchase
1975 Edward M. M. Warburg, B.S., L.H.D.................New York
1969 Joseph W. McGovern, A.B., LL.B., L.H.D., LL.D.....New York
1977 Joseph T. King, A.B., LL.B........................Queens
1974 Joseph C. Indelicato, M.D.........................Brooklyn
1976 Mrs. Helen B. Power, A.B., Litt.D.................Rochester
1979 Francis W. McGinley, B.S., LL.B...................Glen Falls
1981 George W. Weinstein, LL.B.........................Hempstead
1980 Max J. Rubin, LL.B., L.H.D........................New York
1971 Kenneth B. Clark, A.B., M.S., Ph.D................Hastings on Hudson
1982 Stephen K. Bailey, A.B., B.A., M.A., Ph.D., LL.D..Syracuse

President of the University and Commissioner of Education
James E. Allen, Jr.

Deputy Commissioner of Education
Ewald B. Nyquist

Associate Commissioner for Research and Evaluation
Lorne H. Woollatt

Assistant Commissioner for Research and Evaluation
William D. Firman

Director, Division of Research
Carl E. Wedekind

Director, Division of Evaluation
Alan C. Robertson

The Educational Research Association of New York State

OFFICERS—1966

President—Richard M. Clark, State University College at Oneonta
Vice President—S. David Farr, State University of New York at Buffalo
Secretary-Treasurer—Leo D. Doherty, New York State Education Department
Directors—Roscoe C. Brown, Jr., New York University
Paul M. Halverson, Syracuse University
Richard E. Ripple, Cornell University
John Rosenbach, State University of New York at Albany

OFFICERS-ELECT, 1967

President—S. David Farr, State University of New York at Buffalo
Vice President—Roscoe C. Brown, Jr., New York University
Secretary-Treasurer—Leo D. Doherty, New York State Education Department
Directors—Paul M. Halverson, Syracuse University
Richard E. Ripple, Cornell University
William C. Rock, Rochester City School District
John Rosenbach, State University of New York at Albany

COMMITTEES

Program

Howard Berkowitz, Chairman, State University College at Oneonta
Leo D. Doherty, New York State Education Department
Howard Kight, State University of New York at Buffalo
Richard Lonsdale, New York University
David Payne, Syracuse University
William C. Rock, Rochester City School District

Arrangements

Leo D. Doherty, New York State Education Department
George T. Whitcomb, New York State Education Department
Maria Horvath, New York State Education Department
Donald Donley, State University of New York at Albany

Speaker Arrangements

Lorne H. Woollatt, New York State Education Department

Publications

Leo D. Doherty, New York State Education Department
Nicholas T. Whelden, New York State Education Department
Elizabeth B. Baker, New York State Education Department

Newsletter

Editors—Julian Roberts, Yeshiva University and
Maurice Lohman, City University of New York
Regional Editors—Howard Berkowitz, SUC at Oneonta
Roscoe Brown, Jr., New York University
William T. Callahan, The Educational Council, Long Island
Richard M. Clark, SUNY at Albany
Leo D. Doherty, N.Y. State Education Department
William Greenham, Brighton Public Schools
Seymour Halpern, Syracuse University
Robert Heller, SUNY at Buffalo
Paul Joslen, University of Rochester
Richard E. Ripple, Cornell University
S. David Farr, SUNY at Buffalo

CONVOCATION CHAIRMEN

Mark Beach, Cornell University
Norman Becker, Cornell University
Edward Fipay, State University of New York at Albany
Robert Heller, State University of New York at Buffalo
Edward Henderson, New York University
Howard Kight, State University of New York at Buffalo
Maurice Lohman, City University of New York
Richard Lonsdale, New York University
Frank W. Lutz, New York University
Richard E. Ripple, Cornell University
William C. Rock, Rochester City School District
Aubrey Roden, State University of New York at Buffalo
John Rosenbach, State University of New York at Albany
John Schmitt, Boston College
Bernard Shore, New York University
Richard L. Wing, BOCES #1

FOREWORD

The theme of the Seventh Annual Research Convocation was "The Emerging Research Worker and His Needs." At a time when problems are looming larger than ever and solutions seem sometimes slow in coming, the convocation justly addressed itself to the place of educational research in the larger social setting and, more specifically, to the role that the educational researcher can and will play. Certainly, the field of research in education is filled with hope and rich with promise towards solving the major problems of our day. The evolving role and function of the researcher—his needs, his uses, and the unique contributions that he has to make—are thus becoming increasingly important in the context of modern times. The measured and thoughtful consideration of these issues was the main thrust of the convocation.

The following pages reflect these ongoing concerns. Major presentations focus on such topics as educational research and the accumulation of knowledge, curriculum modifications for the emerging research worker, the emerging research worker and his needs, and the question of whether the educational researcher shall be a computer scientist or a Renaissance man. Other talks consider such matters as research methodology, taxonomy, and the training of educational research workers. Focusing on specific knowledge in the field, reports of current research discuss such areas as complex educational media, programmed learning, school administration, research related to disadvantaged children, concept formation, instructional mode, reading, and the Elementary and Secondary Education Act of 1965. In all of these, the major theme, the primary implication, is the expanding of educational opportunity for all.

Finally, a word of thanks is in order to all those who contributed to the convocation their time, efforts, and skills. Major speakers, symposium chairmen and participants, chairmen of reporting sessions, and the many individuals who presented papers all helped to make the convocation a worthy and valuable experience. Special tribute is due Richard M. Clark, President of the Association, Howard Berkowitz, Program Chairman, and Leo D. Doherty, Coordinator of the convocation. It is with contributions such as these that we may recall past convocations with pride, and look forward with keen anticipation to the convocations of the future.

LORNE H. WOOLLATT

CONTENTS

Page

FOREWORD v

I. MAJOR PRESENTATIONS

Remarks of Welcome, Lorne H. Woollatt, Associate Commissioner for Research and Evaluation, State Education Department 3

Gaining Some Perspective on Educational Research and the Accumulation of Knowledge, David R. Krathwohl, Syracuse University 5

Curriculum Modifications for the Emerging Research Worker, David S. Bushnell, United States Office of Education 11

Presidential Address, Richard M. Clark, State University of New York at Albany 15

The Educational Researcher: Computer Scientist or Renaissance Man? John B. Carroll, Harvard University 18

II. HIGHLIGHTS OF RECENT RESEARCH

*Learning and Development, Aubrey Roden, State University of New York at Buffalo . . . .

*Social Foundations, Robert Gowan, Cornell University . . . . .

Selected Studies in Administration: Recent Research in Three Social Science Dimensions, Lawrence Iannaccone, Frank W. Lutz & Harry J. Hartley, New York University 25

Highlights of Recent Research in Bayesian Statistics, Donald L. Meyer, Syracuse University . . 30

Highlights of Recent Research Curriculum and Teaching, Glenn E. Reeling, Jersey City State College 32

III. SYMPOSIA

The Demonstration of Computer-Aided Instructional Systems and Other Complex Educational Media, Richard L. Wing, BOCES #1, Westchester County, New York, Chairman; Robert E. Taylor, Mt. Kisco School District; Father John Boyle, Fordham University, College of Philosophy and Letters; Walter Goodman, BOCES #1 Westchester County, New York, Participants . . 35

Needed Research in Teacher Education: Theory Based Programs and Teacher Education, Lawrence Castiglioni, City University of New York, Queens College 38

Needed Research in Teacher Education: Sensitivity Training and the Process of Change, Julian Roberts, Yeshiva University 40

*Needed Research in Teacher Education, Maurice Lohman, City University of New York; Louis Heil, Brooklyn College; David Fox, City College of New York; William Reiner, Hunter College .

Characteristics of New England Catholic School Principals

- Characteristics of New England Catholic School Principals, John A. Schmitt, Boston College, Chairman 42
- NECEC and Research in Catholic Education, Reverend Paul F. McHugh, New England Catholic Education Center 42
- The Studies of Catholic School Principals' Characteristics, Reverend Patrick J. O'Neill, Diocese of Fall River, Massachusetts; Thomas Whalen*, Stoughton, Massachusetts Public Schools 43
- The NECEC Principalship Studies and Administrative Theory, George F. Madaus, New England Catholic Education Center 43
- Implications of the Studies for Higher Education, Vincent C. Nuccio, Boston College 44

Taxonomic Research in Organizational Behavior in Educational Systems, Frank W. Lutz, New York University; Mozell Hill, New York University; Paul Cullinan, New York University; Glenn Immegart, The University of Rochester; Seymour Evans, New York University; Harold Robbins, Queens College 46

Programs for Training Educational Researchers, S. David Farr, State University of New York at Buffalo 53

Page

Training for School Research Directors, Louis T. Di Lorenzo, State Education Department . . . 55

The Training of Educational Researchers, Marvin Taylor, Queens College 56

*The Training of Educational Research Workers, Richard Lonsdale, New York University; Esin Kaya, New York University; Donald Meyer, Syracuse University . . . .

A Cooperative Program Between a City School District and a Suburban School District, William C. Rock, Hermon R. Goldberg & Joanne E. Lang, City School District of Rochester; L. William Heinrich, West Irondequoit; James V. Mitchell, The University of Rochester 58

The Elementary and Secondary Education Act of 1965 Title I, Irving Ratchick, State Education Department 63

*The Elementary and Secondary Education Act: A Progress Report, William C. Rock, Rochester City School District; Frank A. Stevens, State Education Department; Mark B. Scurrah, State Education Department; Richard A. McCann, U.S. Office of Education . . . .

IV. RESEARCH REPORTS

The Facilitation of Problem Solving and Verbal Creativity by Exposure to Programmed Instruction, Richard E. Ripple, Cornell University; John S. Dacey, Boston College 69

Relationships of Some Student Characteristics and Verbal Creativity, John S. Dacey, Boston College; Richard E. Ripple, Cornell University 70

The Relationship of Anxiety, Creativity and Intelligence to Success in Learning from Programmed Instruction, Robert P. O'Reilly, Richard E. Ripple & Donald J. Treffinger, Cornell University 71

1,000 Programs of Self-Instruction for All Ages: An On-going National Survey, Jonathan W. Varty, Brooklyn College, City University of New York 74

A Report of An Evaluation of the Teacher Education Program of the School of Education of Long Island University, Richard P. Harmon, Long Island University 77

A Monte Carlo Model for the Prediction of Public School Enrollments, Mary Griffin & John Schmitt, Boston College 79

Performance of School Districts on Fiscal and Other Factors in Relation to Agencies of Budget Approval, Charles M. Bernardo, Central School Boards Committee for Educational Research . . 81

A Study of Teacher Tenure in New York State and Survey of Teacher Tenure Laws in the United States, William J. Hageny, State University College at New Paltz 83

An Evaluation of Title III (NDEA) in New York State, William T. Callahan, The Education Council for School Research and Development 85

A Discriminant Analysis of Variables Used To Select Students from Disadvantaged Backgrounds, Roscoe C. Brown, Jr. & Edward H. Henderson, New York University 86

The Contribution of Selected Variables to the Prediction of Achievement in a Summer Compensatory Education Program for Entering College Students from Disadvantaged Backgrounds, Edward H. Henderson & Roscoe C. Brown, Jr., New York University 88

The Relationship of Rigidity and Authoritarianism to the Achievement of College Students from Disadvantaged Backgrounds, Dorothy Lander, Harry Ploski & Roscoe C. Brown, Jr., New York University 89

Teaching Disadvantaged Negro Children To Read: A Second Report on the CRAFT Project, Albert J. Harris, The City University of New York 91

Sex Differences in the Development of Concepts of Space, Matter and Energy in College Students, Leona K. Adler, City University of New York, Hunter College 92

Cognitive and Intellectual Factors in a Study of Engineering Students, Henry L. Rosett & Herbert Robbins, Department of Psychiatry, The Mount Sinai Hospital; Walter S. Watson, The Cooper Union 94

A Method for Studying Verbal Behavior During Problem Solving, Ralph O. Blackwood & Barbara Newman, The University of Bridgeport 96

Page

Cognitive Development of Children Between Two and Seven Years, Greta Morine & Ruth Formanek, Hofstra University . . . 98

Models for Classification, Robert F. McMorris, State University of New York at Albany . . . 100

Nonlinear Payoff Functions, Kenneth I. Raymond, State University of New York at Buffalo . . 102

On the Use of Interaction Terms in Multiple Regression Equations, Richard B. Darlington, Cornell University; Dieter H. Paulus, The University of Connecticut . . . 103

Item Sampling, Robert Barcikowski & Carmelo Terranova, State University of New York at Buffalo 106

Headstart: Evaluation with a New Instrument, Ruth Formanek, Hofstra University . . . 108

Changes in Children's Scientific Concepts as a Function of Object Interaction, Social Modeling and Attention, John L. Reed, Skidmore College . . . 111

The Interaction of Sex of Teacher and Sex of Pupil as a Factor Affecting Teachers' Rankings of Pupil Anxiety, Donald J. Treffinger & Richard E. Ripple, Cornell University . . . 113

Planning Instructional Materials for Certain Lower Class Elementary School Children, George Movesian, Hunter College, City University of New York . . . 114

Evaluative Study of Prekindergarten Programs for Educationally Disadvantaged Children, Louis T. Di Lorenzo & Ruth Salter, New York State Education Department . . . 115

What do Elementary Students Think of Science Fairs? Raymond G. Kenyon, Mid-Hudson Regional Supplementary Educational Center, New Paltz; William J. Hageny,* State University College at New Paltz . . . 117

Effect of Varied Presentations of Laboratory Exercises Within Programed Materials on Specific Intellectual Factors of Science Problem Solving Behavior, Donald Nasca, State University College at Brockport . . . 119

Predicting Success in Modern Foreign Languages, Harold F. Bligh, Suzanne Farren & Geraldine Von Maluski, Harcourt, Brace & World, Inc. . . . 122

Subjective Analysis of an Experiment Comparing a Team and a Conventional Approach to the High School Biology Course, Paul Joslin & John Montean, The University of Rochester; John Schmitt, Boston College; David B. Robinson, Greece Central Schools . . . 125

Personality and Attitudinal Correlates of Alienation in a High School Population, James V. Mitchell, Jr. & Roger M. Weir, The University of Rochester . . . 127

Validity Concomitants of Cummings World History Test, Otis Quick Scoring Mental Ability Tests: Gamma Test, and New York State Regents Examination in World History, Norman C. Maberly & Harold F. Bligh, Harcourt, Brace & World, Inc. . . . 130

Student Drinking Habits as Related to Personality Variables, S. W. Johnson, Thomas Carlsen, & M. L. Levine, State University College at Plattsburgh . . . 133

Students' Perceptions of College Experiences Related to Dropout and Achievement, Horace B. Reed, Skidmore College . . . 134

Cooper Union OAIS Validity Study—1966, Walter S. Watson, The Cooper Union . . . 136

Four Year Development of College Students' Values, S. W. Johnson, State University College at Plattsburgh . . . 139

Social Distance and Deviancy Among College Students, Louise W. Fox, Finch College; David J. Fox, The City College of New York . . . 140

Critical Review of the Stanford-Binet I.Q. Tables, Conwell Higgins, Albany Public Schools . . . 143

The Relation Between a Disguised and Reported Measure of Risk Taking on Objective Examinations, Malcolm J. Slakter, State University of New York at Buffalo . . . 145

*An Analysis of Some Characteristics of Respondents and Non-Respondents to Two Different Types of Mailed Questionnaires, Seth N. Leibler, The Communicable Disease Center . . . .

Problems in the Assessment of Group Counseling, David G. Zimpfer, The University of Rochester . . . 147

Assessment of Attitude Toward School: A Preliminary Study, Edward E. Panther, State University College at Buffalo . . . 149

Page

Characteristics of Low and High Performing Elementary Pupils, John H. Rosenbach & Robert F. McMorris, State University of New York at Albany; William Smith, Guilderland Central Schools 151

The Major Purposes and Functions of Supervision as Perceived by New Jersey Public School Superintendents, Supervisors, and Building Principals, Geraldine F. Cleminson, Fordham University 152

A Statistical Model of Conformity-Deviance, Joan R. Egner, Cornell University 153

Perceived Need Deficiencies of Teachers and Administrators: A Proposal for Restructuring Teacher Roles, Francis M. Trusty & Thomas J. Sergiovanni, The University of Rochester . . . 155

Personnel Practices in Special Education, Robert E. Mulligan, Public Schools, Fords, New Jersey . 157

Instructional Mode, Personality Characteristics, and Achievement: Overview of the Research and Development of Instructional Materials, Richard E. Ripple, Cornell University 159

Development of the Criterion Test Learning Structures Project, Jason Millman & M. D. Glock, Cornell University 161

Instructional Mode, Anxiety, Exhibitionism and Achievement, Robert P. O'Reilly, Cornell University 163

Instructional Mode, Compulsivity and Achievement, Lawrence Wightman, Cornell University . 166

Instructional Mode, Achievement and Creativity, John S. Dacey, Boston College 168

Evaluation of Reading Progress of Fourth Grade Pupils Enrolled in the Remedial Reading Program of the After School Study Centers, George Forlano & Jack Abramson, New York City Public Schools 169

Beginning Reading—The Effectiveness of i.t.a. and T.O., Harold J. Tanyzer, Harvey Alpert, & Lenore Sandel, Hofstra University 171

An Evaluation of the Initial Teaching Alphabet in the First Grades of the Canandaigua Elementary School, Marion Walsh, Canandaigua Public Schools 173

55

Relationship of Reading Achievement and Motivation, Irving Zweibelson, New Rochelle Public Schools; Francis J. Lodato, Manhattan College 175

V. LIST OF PARTICIPANTS 177

---

*Papers presented at the Convocation but not presented for publication by the speakers

Major Presentations

REMARKS OF WELCOME

Mr. Chairman, Members of the Convocation:

Welcome to the Seventh Research Convocation jointly sponsored by the Educational Research Association of New York State and the New York State Education Department.

In order to appreciate the progress we have made, you would need to have attended the annual meeting of the Association in 1959. A handful of people met in Syracuse after the conclusion of the Conference of the Secondary School Principals Association. Business consisted simply of an apology for lack of progress in 1958-59, the election of officers, and a promise to do better.

In the ensuing months, the Education Department agreed to sponsor a joint research convocation in 1960. This was a turning point in the organization of educational research in New York State. Today and tomorrow we expect 400 people to attend. As many as 11 states and two Canadian provinces are likely to be represented. A group of undergraduate students exploring educational research is with us.

The growing number of people in attendance is the result of the worthwhile nature of the program. Major speeches will relate to the theme of "The Emerging Research Worker and His Needs." Dean Krathwohl of Syracuse University will start the theme in a few minutes dealing with "Perspectives on Educational Research." After lunch, Dave Bushnell's title will relate to "New Curriculum Modifications for the Emerging Research Worker." At this evening's banquet John Carroll of Harvard, in his inimitable style, will talk of "The Educational Research Worker: Computer Scientist or Renaissance Man?"

One of the most significant developments is that there are 25 educational research trainees at this conference. They are the first wave of participants in New York State's Training Program for Educational Research Workers. The program is cooperative; the major schools of education in the State accepting several trainees each. It is also cooperative in the sense that funding results from a special grant to the State Education Department under Title IV of the Elementary and Secondary Education Act of 1965. A second wave of 25 trainees will enroll in 1967. A further increase of 50 is anticipated later. Trainees will study research on college campuses for a year. They will spend the second year in supervised field experiences in research in education. By June 1968, a number will be available as research workers in school systems. If you want more details, attend Symposium Number 5 at 11:00 a.m., Tuesday.

The program is not limited by the theme. Eleven reporting sessions and symposia will range through research in programmed learning, psychology, instruction; methodology and computer aided instruction will be discussed. The closing session will deal with the new massive infusion of Federal program funds.

And so I wish you a professionally happy two days. Get immersed in research. Enjoy our mutual fellowship and bring a friend in 1967.

LORNE H. WOOLLATT

GAINING SOME PERSPECTIVE ON EDUCATIONAL RESEARCH AND THE ACCUMULATION OF KNOWLEDGE

David Krathwohl, Syracuse University

The now rapidly developing changes which are taking place in educational research are signs of a healthy and growing science. To maintain this forward movement we need to understand where we are and, having gained this perspective, how we may move further.

Comte noted over a century ago that *each of our leading conceptions,–each branch of our knowledge, –passes successively through three different theoretical conditions: the theological or fictitious; the metaphysical, or abstract; the scientific, or positive.* Education has been for centuries, and still is to this day, the subject of much "armchairing" and abstract conceptualizing. It is only recently, however, that we have begun to achieve the third stage—that we are skilled enough in our powers of observation, meticulous enough in our documentation, and clear enough in our logic to be considered by others, and to begin to accept ourselves as "scientific." Empirical educational research is in its infancy.

Although excellent research texts were available thirty years ago—texts such as Monroe and Engleheart—relatively little research based on the methods they advocate was undertaken until the advent of the Cooperative Research Program in 1957. This program, administered by the U.S. Office of Education, began with a million dollars and has grown now to 70 million. A 70-fold growth in funding in a decade should be a matter of encouragement, but looking at the Federal pie as a whole, this slice is still woefully small. Recently, when the defense budget was some 70 billion, expenditures for research and development comprised about 7 billion—10 percent—of the total.

The health and medical field is estimated to be a 38 billion dollar industry in which about 2 billion is devoted to research and development, somewhat over 5 percent. Educational expenditures probably amount to about 45 billion and yet a bare 150 million, about one-third of one percent, finds its way into research and development. Clearly, we are far from being in the same research league financially with other areas which vitally affect our society.

As might be expected of an infant science, our research projects tend to be small in a statistical sense, as well as dispersed and unconnected in a systems sense. We have few large studies such as *Project Talent* and the research recently completed by the U.S. Office of Education, the *Equality of Education, Educational Opportunity Study*. The rule has been studies at the 30 to 60 student level.

What do our present activities and our methods of research add up to? Inexperienced as our science is, we are making some important inroads on significant problems. We know better now than before, for example, how to analyze the role of the teacher, to determine by objective observation and assessment, the effectiveness of teaching in terms of its differential effects on student performance.

Nietzsche wrote that "in large states public education will always be bad, for the same reason that in large kitchens the cooking is usually bad." But, within the last few years systematic research has seen marked gains in our ability to devise better ways of individualizing instruction.

We've begun to see the role of the administrator much more clearly, to understand his relations with his board, how he chooses and relates to his teachers, how he engages the community in solving the problems of education, especially in relation to that bitter pill, the bond issue.

One important advance of our recent research has been an awakening to many ways in which new ideas are disseminated and adopted and the criticalness of more knowledge of this kind to change the schools.

I could list other evidences of progress and, when I was through, you could add to it. The point to be made is that we have been gaining knowledge about the educational process that can make a difference. Assuming this, two questions seem paramount. First, are we progressing as rapidly as possible? Second, are we structuring our research so as to facilitate that progress? We need to be concerned with how to increase the speed as well as the effectiveness with which we accumulate knowledge. It is on these matters that I will suggest 9 points of varying importance. Let us run through them quickly to set the scene and then return to each for a more detailed look.

First, the logic of hypothesis testing should put the emphasis on the certainty with which knowledge is held;

Second, more meaningful measuring scales should increase the practical significance of new knowledge and facilitate the study of the inter-relationship of variables;

Third, study of the effect of a variable throughout its range of values should be more conducive to theory building than merely knowing the direction of effect (implicit here is the important assumption that theory building is critical to knowledge accumulation);

Fourth, and perhaps heretically, frank disclosures of the shortcomings of studies should greatly ease the determination of their contributions and facilitate the design of complementary studies;

Fifth, a study of the seriousness of certain violations of accepted practice might relieve us of restrictions on our research;

Sixth, a re-examination of present criteria for the dissemination of the results of research to the academic community might suggest improvements important to the accumulation of knowledge;

Seventh, replication of studies should be increased as a necessary part of the validation process;

Eighth, shortening the time span between inventor, innovator and discoverer, and the consumer, can be achieved by more convincing design of studies, by more and more careful and explicit exposition of the relationship of the immediacy of these studies to the consumer's notions of the real world, and by the exploitation of every sound technique of information retrieval and dissemination;

Ninth, research resources should be spent where the largest payoff at any stage of development of a research area can be expected.

Now, let us examine these points more closely:

Improving the Logic of Hypothesis Testing

I have suggested to you that the logic of hypothesis testing needs to be refocused. Our present approach involves the disconfirmation of a null hypothesis—rather than a positive validation of those ideas which merit acceptance. We are seeking a binding identity between an hypothesis and that fragment of the real world it attempts to describe.

The disconfirmation technique gives us no way of stating how much more certain we are that the principle or knowledge being tested is true at the end of the study than we were at the beginning. Each study is merely one more link in a chain of evidence upon which a judgment must be made. If a study is a tight one, and the expected effect does not occur, a total invalidation of the hypothesis is achieved. But, a set of results which confirm provides an indication, rather than a total proof of soundness. This is the logic of the situation, and it is only as an hypothesis repeatedly escapes disconfirmation in a series of tests that it can be accepted, and then only conditionally.

The essential value of Bayesian statistics is that it focuses on the certainty with which knowledge is held true. The researcher is required to evaluate the accumulated evidence in support of his proposition. From this, he establishes in probability terms his first-level confidence in the proposition. Then, continuing his investigation by experiment, this initial *a priori* probability is modified to reflect appropriately the fact that evidence disconfirms or confirms the hypothesis in question and, thus, a second-level probability position is taken. This process, which is carried on to the limits of practicality, has the advantage of developing a statement of where the researcher stands in the confirmation process, as well as the increment or decrement in probability terms that each experiment adds.

The Bayesian approach emphasizes understanding of the sureness with which knowledge is held, and thus permits the application of this knowledge within a set of reasonable and realistic expectations. The present limitation on the technique is the problem of establishing a reliable process for determining *a priori* probabilities. Progress, however, is being made and enough is already known to use the Bayesian statistical method on a number of problems. It appears reasonable to expect that as we widen the application we will greatly speed up the accumulation of knowledge and strengthen our ability to use it.

The Need for More Meaningful Measuring Scales

My second point suggests that we can facilitate the accumulation of knowledge by developing more meaningful measuring scales. The most useful that we now have are those tied to normative bases such as age or grade norms. Typically, however, the reader of educational research is faced with, let's say, a two-point gain on an unknown test and the assurance of the experimenters that the gain is significant at the 5 percent level of confidence. When the test is especially constructed for the experiment, the reader is usually left badly adrift. What do these results mean to the decisions that must be made? How much is a two-point gain worth in dollars as an improvement in learning methods? Does the 5 percent level assign the principles developed in the experiment to the category of unacceptable risk?

If we could tie more meaningful tags to the gain scores, we might achieve the practical end of saying that a two-point gain will net the student "x" dollars in increased salary when he graduates or will result in "y" minutes in time saved in bringing a student to a given level of achievement in school. It is admittedly difficult, and unquestionably dangerous and distasteful, to try to assign monetary values to learning, since it requires putting a price tag on intellectual and humanistic values. Still, we need to find some way of bringing meaning to our measuring scales. From the point of view of the practicing educator, the use of monetary or time standards may be a welcome first step. We seem to have largely abandoned research in this area, but it is one that would pay very large dividends if better solutions could be found.

The Inter-relation of Variables

The third proposition concerns the improvement of our understanding of and ability to work with the inter-relation of variables. We have developed analysis of variance techniques to a level of sophistication which is sufficient to analyze very complex problems. But, analysis of variance tells us only the direction of effects and which of the effects cannot be charged to error. Unless a study is so designed that the experimental variable is present at several levels of strength, it does not permit the experimenter to plot the way one variable affects one or more of the others as it changes in intensity. For example, most recent studies of achievement in schools where schoolrooms have been integrated across social class lines indicate advantages in favor of the lower class children. But these studies generally lack persuasive force because the existence of incremental variations, so important to demonstrating generalizing principles, has not been investigated. One might ask: How does the proportionate mix in social class relate to the average achievement score? Is it maximized when lower class children are in the minority? In the majority? When the classes are equal? At some point of optimization? It seems clear that a plot of the function relating mix to achievement scores and further delineated by the impress of other related variables would be most useful in understanding how these functions interact.

Combined with the plea for more meaningful metrics, this further suggests that we would rediscover the neglected art of curve fitting. This process was regularly taught in statistical courses in the 20's but has lapsed into disuse. Curve fitting gives precise information for theory building that is available no other way. It may be time to reconsider its use.

The Compromise in Research Design and Its Disclosure

The fourth area of importance is taking conscious action to make clear the weaknesses as well as the strengths of our studies. Doctoral dissertations typically include a section called "Limitations of the Research," intended to help the candidate circumscribe his problem into reasonable limits. But the necessity for doing a study within reasonable limits of time, energy, finances, and available situations is hardly confined to doctoral candidates. We all face it. In a very real sense, nearly every study is a compromise with ideal criteria. The complexities of the phenomena we study and the fact that the criteria for judging practical studies are not always internally compatible, force such compromises.

The researcher who gathers his evidence from a real situation, such as the classroom, is faced with a host of factors to control if he is to reason his way to a connection between observed effect and experimental variable. In a teaching methods experiment, for example, he must control for differences in the abilities of teachers, for differences in their enthusiasm for one method or another, for Hawthorne effect, for pretest effect, for pupil abilities, for socioeconomic environmental factors, for academic climate and school environment, and so on. If this confusion drives him to do his research in the laboratory, using neutral material such as nonsense syllables and memory drums, he may in the end have compounded his dilemma. Now, he is faced with problems of relating his results to the classroom and an uncertainty of whether and how they will generalize, which can be resolved only by actually doing the study he tried to avoid.

If we can get each investigator (who probably knows the nature of his compromise better than anybody else) to describe it and the rationale for this choice in his report, if we can convince editors and their advisory staffs that such a study is increased in value by such candor, readers will find it somewhat easier to judge the value of the study for themselves, other researchers could more easily design complementary studies with different and offsetting compromises.[1]

I can further illustrate the complexities of choice leading to compromise by a recent case involving a study of programmed material. Most such studies have been done with relatively few cases. Thus, the common finding of "no significant difference" may result from the fact that there are insufficient cases to provide a precision adequate to sense the experimental effect. Now how should one increase the size of the group yet keep experimental control? Ideally, the program would be administered in a teaching machine, so we would know the student went through the program step by step, didn't peek at the answers, had a quiet room, and so on. On the other hand, to do so in the context of the course would require a great many machines, suitable space, supervision, and tight scheduling, if the pace of the course were to be maintained. Finally of course, one needs the money to pay for all this. In this particular instance, the compromise was the use of a large sample of students who took the materials home with them to study, thus relaxing some control over how they would use them. The increase in sample size markedly increased the sensitivity of the study but, of course, at the cost of some loss of control. At the same time, the conditions used were closer to those under which the materials would actually be used in practice and so increased the practical generality of the findings. This is an example of the kind of compromise that faces every investigator.

---

[1] Alternatively, a journal critiquing current periodical literature would probably better fit the current publishing pattern and might at least partially serve the function described. Similarly, more critical reviews of research could facilitate this purpose.

The Evaluation of Assumptions Underlying Research

The fifth suggestion on my list is closely related to the fourth. As noted earlier, it is rare that a study does not violate at least one of the criteria by which its quality would be judged. Yet, we have only some beginning studies to tell us how serious such violations are.

For example, many of us have long considered the Hawthorne Effect a serious problem unless properly controlled. Hawthorne Effect results from the special attention given to an experimental group, which may itself affect the dependent variable even though the independent variable does not. The concept developed in the early twenties when Roethlisberger, Dixon and others conducted experimental investigations of worker motivation at the Hawthorne plant of the Western Electric Company. Now, some 40 years later, Desmond Cook's study of this phenomena raises some very serious questions about the extent to which this effect is as pervasive and strong as previously supposed.

The violation of the assumptions of statistical tests is a common breach of criteria, sometimes considered serious enough to invalidate the inferences drawn from a study. But a study such as Norton's, as reported in Lindquist's *Design and Analysis of Experiments,* concerning the "F" test and similar studies on other tests have shown these tests to be astonishingly robust. Even flagrant violations do not produce serious effects which cannot be compensated for by a simple correction.

We have had far too few such studies as Cook's and Norton's. It will no doubt pay us to regularly devote some portion of our resources to research on research itself, with a special emphasis on a re-examination of some of our dogma.

The Need for Better Dissemination Techniques for New Knowledge

Sixth in the chain of concerns is the variety of gatekeepers on the road between the generation of new knowledge and its acceptance. These are such persons as the administrators and panel members of funding agencies, the journal editors, and the textbook authors and publishers. Their points of view may facilitate or may impede the accumulation of knowledge. There are two fairly common policies which appear on the surface to inhibit knowledge accumulation. One of these is the heavy emphasis on the part of the editorial establishment favoring empirical research over various kinds of theory building or "think pieces." Skinner's article in Science heavily influenced the programmed learning movement, but there are few comparable efforts today appearing in the journals. Editorial policy may be only part of the cause, for such articles are less likely to be prepared if publication outlets are not available. It is the familiar chicken and egg situation, with such articles likely to stimulate additional such production.

The second policy is the burial without notice of research which produces negative results. Perhaps our gatekeepers are printing all such studies that are worth recording, but this is questionable. Let me document this point from personal experience. We were recently thinking about designing a program to attract women, who had been teachers or taken teacher training, back into the field after their family responsibilities lightened. We were interested in the problems of remotivation and retraining. We understood that several programs attempting this same task in the post-World War II years had not been successful, but the difficulties of such programs did not appear in the usual publication channels. This makes it very difficult to avoid repeating the errors of the past.

We are caught in an information selection, processing and dissemination situation—one hardly can grace it by calling it a system—which, like Topsy, just grew. The above represent two policies which appear to be weaknesses in this system, but we need some studies of it which better indicate how it facilitates and how it impedes the accumulation of knowledge.

Increased Replication To Assure Validity

The seventh matter listed refers to the importance of replication as a necessary part of the validation process. It may seem unnecessary to comment on a point so self-evident, but, for a variety of reasons, we had had little replication in education research. There has been a need to stretch our research funds at the cost of thoroughness. There have been the temptations of unexplored fields and the natural preference of research investigators to adventure into these new fields rather than to secure those already won.

The sad reality of research, especially in the social sciences, is that replication is one of the few techniques of validation available to us. If our new knowledge is to find its way into the mainstream of education and if educational science is to mature, we must engage in replication and endure its burdens of cost and time.

Shortening the Time Between Discovery and Use

The eighth point concerns the very long time from innovation or discovery to the consumers' awareness that new knowledge is available. A recent study by the American Psychological Association found that an average of 5 years elapses between the start of a study and its citation in a research review, and of these 5, at least 3 years elapse to journal publication. In addition to these years, time required to write a funding proposal, wait for its approval, and obtain actual funding can easily add 6 months, a year, or more. Finally, of course, if the study has practical implications, there is the considerable period of time involved in getting the results accepted and implemented in the field.

A number of efforts are under way to reduce this time lag, and studies have begun of the dissemination and adoption process. Hopefully, ERIC, the Education Resources Information Centers of the U.S. Office of Education, will reduce proposal preparation time by helping researchers retrieve relevant research.

Perhaps the biggest lag is in the use of research results. At one time, the delay in adoption of new practices was estimated to be as much as 40 years. This situation appears to be improving. The new mathematics programs, for example, seem to have been widely adopted in about a decade.

Could researchers reduce field dissemination time further by building more convincing studies? Investigators are concerned primarily with communicating with fellow researchers, and they build studies primarily in the form and to the size needed to achieve statistical significance. What may be relatively convincing to a fellow researcher, however, may not be even minimally convincing to a research user. In developing certain studies, the researcher needs to cast a glance over his shoulder at the waiting, potential users, and build his case from the start which, if the expected results appear, will be convincing. Small, carefully built, carefully controlled studies contribute knowledge and assist in building a convincing impression, but sample size and number of replications are particularly important.

The Equality of Educational Opportunity Study of the National Center for Educational Statistics is an example. This was an attempt to document the conditions under which Negroes attend school. A small study would have failed in its task. This study makes its case by involving 4,000 schools, 20,000 teachers, and 645,000 students. Implications and principles of action for the bulk of the Negro community are much more certain with sample sizes of this magnitude. Educational research is only beginning to develop studies of this complexity and magnitude. We ought to consider with some care whether there are not more instances where the results of a study would be greatly strengthened, and the transfer and use of new knowledge greatly accelerated by such augmentation of the data base.

It will be helpful too, if we do a better job of helping the consumer understand the method of science. We say that we believe that knowledge is only held as true until it is replaced by new knowledge. But the continuum on which a scientist scales the certainty of the knowledge he holds is not usually made obvious to the consumer, who acts as though the choice is dichotomous, either true or false. Better understanding of the uncertainty of findings might open consumers to a willingness to change, to understand the occasional starts, stops, and perhaps even reverses in the accumulative process. For example, the lack of empirical support for the original principles used to shape programmed instruction is less confusing to the researchers, who understand how the process could go awry, than it is to the research consumer, who takes as gospel the programming principles originally developed.

There is something of a parallel to be drawn with the fourth point which concerned the need for disclosure of the limitations of research projects. Such limitations and uncertainties, properly explained to the consumer, could bring about greater understanding of the scientific process and hasten rather than impede the application of knowledge.

Such publications, designed for nonspecialists, as

Washington University's *Transactions,* which explains social science research, and *Scientific American* which explains a variety of science areas, have helped reduce the communication lag between researcher and consumer, and give some perspective on the research process. Perhaps such a journal for education would be something that we ought to consider.

Choosing Appropriate Targets for Research Effort

The ninth and final area for comment deals with the dilemma of appropriately using our resources for maximum return. Some of the same forces that have acted to discourage replication have also tended to spread our research over a variety of areas. Thus, in the past, research funds in education have, in general, been relatively unconcentrated. But Congress in particular, the public in general, and increasingly of late, the U.S. Office of Education, have assumed that our problems can best be solved by important concentrations of funds and effort. In general, however, concentration of resources is more likely to result in funding leads that do not have a return. More blind alleys will be explored as greater risks are taken in attempts to find a breakthrough. Still, such concentrations may well result in solutions. Whose crystal ball is clear enough to gainsay this? In an open invitation research program, researchers typically propose problems they consider ready for solution. This represents their best estimate of where to spend the funds, for they rarely want to explore an area in which they do not expect to succeed. But their choices are often not in the areas of greatest societal need.

It is important that we learn to recognize the concurrence between the desirable concentration of funds and the readiness of a particular area for a breakthrough if we are to gain the best use of our resources. This concurrence may be affected by answers to such questions as: What is the state of the art and the knowledge of research in an area proposed for concentrated funding? (What kinds of funds and efforts will it take to engineer developments in an area? What has been the past success in engineering similar efforts?) What are the alternative uses of funds? (The problem areas ready for concentration of funds may be comparatively trivial in importance.) What is the level of funding available? (A balance is needed between the funds required to engineer the development of an area and the funds available for research in other important areas.)

What is the disposition of the research community toward an area? (Will one be able to command talent that is both motivated and adequately skilled to do the development task—or is this a low prestige area, or an area for some other reason outside the boundaries of a scientific orientation?) Depending on the closeness to solution of a problem, the state of the art of research, the level of funding available especially with reference to alternative ways of spending funds, and the orientation of the research community, there is probably an optimum solution to the employment of our research means. We need to devote some study to the process of resource allocation. We should, for example, check the accuracy of predictions made about specific decisions which attempted to balance variables such as those noted above. In this way we can learn to approach much more rapidly the maximum accumulation and use of knowledge.

Summing Up

We have ranged over a number of areas in which some effort and/or study might improve educational research, either in a limited way or perhaps by an order of magnitude. Summing up, we have considered:

First, a more appropriate statistical logic for investigation;

Second, the need for more meaningful measures;

Third, the increased use of techniques to study the relation between levels of strength of variables and effects;

Fourth, the compromise in research design and the value in its disclosure;

Fifth, a call for a determination of the gravity of violating certain criteria established for "good" studies;

Sixth, the need for re-examining the criteria used by the gatekeepers of research communications;

Seventh, the need for greater replication in research;

Eighth, ways of shortening the time gap between the development of knowledge in research and its transfer in usable form to the consumer; and

Ninth, the need to channel funding into patterns and paths to provide the greatest overall return for education.

These are by no means all the points that one could raise about this topic. But this is a side of research that is seldom given consideration. We are so busy doing research and attempting to apply research findings that we do not examine the process itself. Perhaps this discussion will stimulate you to stand back and similarly try to put the process in perspective. Such efforts are important if educational research is to reach maturity.

CURRICULUM MODIFICATIONS FOR THE EMERGING RESEARCH WORKER

David Bushnell, United States Office of Education

I'm very happy to share with you some of my own views on what's been happening in the Bureau of Research in the U.S. Office of Education. We have approximately $90 million to spend this year. Let me add, however, that this is the year of austerity. We lost 20 million dollars that we had expected to be appropriated. Ten million of it came out of the Elementary and Secondary Education Act, Title IV, reducing that appropriation from 80 to 70 million dollars. The second, and even harder to accept reduction was in my own program, the vocational education research monies, where we were cut by 50 percent from what we had hoped to be a 20 million dollar appropriation to 10 million dollars. This has led us to some agonizing reappraisal of what we have funded. We have continuation costs of about 8 million dollars, and we had already obligated something like 5 million dollars, putting our budget at a 13 million dollar level. Now we have to cut back to the 10 million appropriation. We're not going to be entertaining very many new projects this year. Those that are well designed and worthwhile will have to be carried over to the next fiscal year. I'm not speaking for the entire Bureau of Research, just the monies appropriated under the Vocational Education Act of 1963. Some of the studies that are underway with fiscal '67 funds will have to be reduced. This period of poverty reminds me of a gentleman that we ran across in the hallways of the office the other day who was wandering around, looking a little disheveled and unkempt. He was saying that he had been reading about the President's declaration of war on poverty and he was here to surrender.

You'll be happy to know that we've introduced systematic planning into the Bureau of Research. We have tried to set forth for this year several priority projects that we'd like to fund. Some of these represent the legislative requirements, the regional laboratories, and the research training program, and others, we think, ought to be focused on major problems or new technological developments that offer some hope of overcoming many of the problems being faced by the public and private schools in this country. In the setting up of these priorities, several Division Directors in the Bureau and Dr. Bright got together to talk about how we could better coordinate the activities of each research division. As you know, we now are structured with five divisions, three of which are designed to serve as the R & D arm for the three major bureaus in the Office of Education, that is, the Division of Elementary and Secondary Research, Division of Higher Education Research, and the Division of Adult and Vocational Research. We've been in the process of pinpointing what we mean by these priority areas but we are not quite ready to release them to the research community. I did think it would be appropriate, though, for me to say a word or two about how we are structuring these priorities and perhaps, by way of example, describe in a little more detail one such effort. Several staff members in the Bureau have been assigned responsibility for a given priority. One staff member is to serve as a team captain, drawing upon not only the personnel and resources within the Office of Education to help structure a position paper, but also upon expertise in the field, consultants whom we can bring in as advisory members of the Task Force on a given priority topic. For example, if we were concerned with the problem of improving the administrative pattern for a statewide system of community colleges, we might well look to people such as yourselves for help in planning how to invest research funds in this particular problem area. Once a paper is developed, we are then in a position to release it to the field and solicit proposals which meet the criteria of the position paper. In other words, we're now in the process of emphasizing sponsored research. We will still accept unsolicited proposals, but most funded projects will releate to the particular priority areas that we are committed to this year and next.

In addition to supporting sponsored and unsolicited proposals, current legislation directs us to administer other programs within the Bureau of Research. I'm now referring to such programs as the regional laboratories and to our research training effort that has just been launched. There are some nineteen regional laboratories now in operation or on the drawing board.

About 18 to 22 million dollars is to be allocated out of this fiscal year's money in order to support the implementation of these laboratories. Their primary purpose, of course, is to make information and consultant help available to local school districts, to encourage the implementation of really significant and promising innovations. The laboratories are designed to bring together researchers and practitioners from the universities and local school districts. In many parts of the country, groups that are now cooperating in the regional education research laboratories had not had much communication with one another previously. Overall, the Office of Education last year spent about 10 million dollars on this program and we have budgeted 18 to 22 million dollars this fiscal year. Dr. Bright sees the new laboratories as the necessary link in the chain of progress that leads from basic research to widespread classroom practices.

The research training program is an effort in which I hope all of you have considerable interest. It began in 1965 with the passage of the Elementary and Secondary Education Act. Eight million dollars were allocated to it last year, and we expect to fund it at about that level again this year. Its purpose, of course, is to help train a larger number of educational researchers in this country than have been trained in the past. There are only about six thousand people now engaged in educational research. We expect that by 1972 there ought to be well over one hundred thousand in this area. The purpose of the program is to improve the quality of educational research training, and to increase the number of training opportunities by expanding existing programs and adding new ones, as well as by providing support to trainees and enabling more students to prepare for careers as educational researchers. You will be interested to hear that the State of New York received over a million dollars last year to implement such programs. One of the more interesting grants, and really the largest single dollar appropriation, some $323,000, was awarded to the New York State Education Department. The Department, in cooperation with several universities within the state, developed a program designed to meet certification requirements for appointment to educational research positions. Fifty trainee positions were authorized under this initial funding. There are other research training grants at Columbia, Cornell, Fordham, Syracuse University, and even one to Rockland Community College, for the support of a research training institute for junior college personnel with research interests. Also, I don't know whether you've seen the recent announcement that we've just published on a new National postdoctoral fellowship program to be awarded by the U.S. Office of Education. If you're interested, you can write a note to Lee Burchinal, Director of the Division of Research Training and Dissemination at the Office of Education, and he'll be happy to send you all the details on this effort.

The National post-doctorate fellowship program is designed to provide opportunities for a limited number of fellows to engage in an intensive postdoctoral training experience with some of the Nation's outstanding researchers. In other words, these fellowships will be allocated to R & D Centers, regional laboratories, or other institutions which are eligible to receive funds under the education research training program. The fellowship stipend would be equivalent to the salary such a recipient would have received through continuation of his regular employment. In addition, there is a $2,000 allocation for travel and incidental cost. Recipients will be invited to attend an invitational seminar conducted at the U.S. Office of Education at the opening of the program, and also a second invitational seminar at the conclusion of one year's residence in the program.

Returning to the topic of research sponsored by the Office of Education, I would like to present to you a concept which we have recently adopted; that of sponsored versus unsolicited research projects. As an example, I would like to describe one example of a systematic attack on a major problem facing American education so that you might better understand what we mean by a program of sponsored research. Several months ago, we became concerned that the 20 or so million research dollars invested in curriculum development projects were not yielding a successful return on that investment. Too many of the projects were small scale efforts with little hope of their replication elsewhere. Consequently, we established a systematic plan for revamping the high school and post secondary or junior college curriculum to better serve the needs of both college bound and noncollege bound youngsters.

Unfortunately, much of what now is being taught in our public schools fails to recognize that technology has generated profound changes in the nature of work. The tendency to separate those in the college preparatory programs from those in the vocationally oriented programs has handicapped both groups. Those who plan to go on to college are not prepared to cope with the question of what happens if they have to drop out—and as you know a great number of students entering college do drop out before completion of the bachelor's degree. Those in the vocational programs find themselves being trained for a narrow range of job skills and receive too little positive reinforcement to develop confidence in their basic learning skills. They not only need to have the kind of entry level qualifications that get them the job in the first place, but they must have the kind of cognitive skills and problem solving capabilities that will allow them to progress in that job or to qualify for new careers. Eight out of ten students in our classrooms today at the secondary level will wind up in jobs requiring less than a bachelor's degree.

If we were to employ vocational preparation as the principal motive, it's our feeling that the inculcation of some of the basic learning skills, reading, writing, mathematics, could be made more palatable to many students.

We feel that a massive research effort is needed to develop a curriculum which would provide every student with the job skills he needs to enter into his first job and to continue his employment even if he were to drop out before completion of high school. The student who graduates from the program should possess the necessary qualifications for maximum flexibility in his post high school activities. He might enter a university or a community college and pursue an academic program. He might enter a community college or a technical school and receive post high school occupational training. He should also have entry-level occupational skills which permit him to go to work. He should have the additional option of continuing his education in an adult education program. The key point is that he should have several options available to him after graduation from high school.

An Organic Curriculum

The first step in building such a student-centered curriculum is to study those behavioral attainments needed by the individual for entry into a variety of post high school activities. Whenever possible, these requirements should be stated specifically and in measurable behavioral terms. Following the lead of the systems analyst, we should describe as specifically and precisely as possible the learning experience which would lead to the desired behavioral outcomes.

The ingredients of a high school program which will assure the attainment of these specifications will certainly include academic as well as occupational training but may also include such components as personal development, real work experience, and individualized vocational counseling. Even the avocational or school-sponsored recreation or social programs may be considered an integral component in this system. Each of these components and subparts must be defined in terms of its contribution to the attainment of the specified behavioral objectives.

The most important feature of such a curriculum is that it is learner oriented rather than process or subject matter centered. The integration and interaction of the components will be a result of careful systems design. There will be no discrete demarcation between academic and vocational skill training or between these and other parts of the system. The truly integral curriculum must be developed so that each activity relates logically to all other activities and leads to the efficient attainment of the behavioral goals.

Such a curriculum is presently feasible and can produce significant improvements in the learning process. An "organic" curriculum would necessarily have to be interesting, challenging, and motivating to each student. It would probably utilize appropriate self-paced and self-instructional technology and maximally accommodate individual differences in learning rate. It should be designed so each student will succeed and yet it should be rigorous in level and content. Furthermore, after thorough experimentation and revision, the integral curriculum should be capable of implementation in, or adaptation to, many different comprehensive school systems in the Nation; and it should be cost effective in the implementation stages.

Summary of Various Elements of This Grand Design

First, we hope that this curriculum effort will emphasize the articulation between the academic and the vocational for the purpose of fusing the two programs. Employing the vocational preparation as the major vehicle, or motivation tool, the inculcation of basic learning skills could, we think, be made more tolerable to many students who have a hard time seeing the value of general education.

Second, we would like to expose students to an understanding of the real world through a series of experiences which capitalize on the universal desire of youth to investigate for themselves. Abstract and verbal principles could be acquired through nonverbal stimuli such as seeing and feeling and interacting.

Third, we would like to develop a core of skills which relate to a cluster of occupations rather than just a specialized manipulative skill in one occupational area. Students with that type of background could, if necessary, transfer to a related type of occupation without a great deal of retraining.

Fourth, we would like to orient students through this new curriculum to the attitudes and habits which go with successful job performance. It's quite obvious to many employers that most workers are not able to perform the jobs assigned to them. Those who get fired are the ones who, because of attitudes, are unable to accept responsibility for themselves; simple things like getting to work on time, getting along with their peers, and accepting supervision.

Fifth, we would like to provide a background for the worker by helping him to understand how he will fit into the economic and civic institutions of his community. There is now considerable work underway at Ohio State, at Columbia, and elsewhere, on developing a better way of presenting the concept of a free enterprise economy to students at the junior and senior high school levels so that they might gain some understanding of the economic and social structures within which they will operate before going on to prepare for whatever occupation they may choose, be it at the professional or subprofessional level.

Sixth, we want to maximize the opportunity for a student to develop and enjoy a positive learning experience regardless of his level of ability and background. A lot of work is being done, as you well know, on programmed instruction and individualization of the instructional process in such a manner that a student can derive a positive reinforcement as he progresses through the curriculum materials.

Seventh, we would like to help students cope with a changing labor market through developing their problem solving abilities and a series of career strategies which lead to an adequate level of income and responsibility.

Last, and perhaps most important, we would like to create within the student a sense of self-reliance and awareness which leads him to seek an appropriate career and aspirational level.

As you see, we're really not talking about revising just vocational education. We're talking about a much broader program to modify curriculum for the elementary and secondary schools. We hope through team teaching arrangements to involve teachers from various disciplines, bringing them together to work out an integrated (with a small "i") curriculum for these noncollege bound students or those who are likely to leave school at some point before they complete their baccalaureate degrees. In general, this type of effort should tap the total capacities of the individual, including his intellectual, manipulative, creative, and social capacities, as well as emphasize the discovery of abstract concepts through the examination of real world events and learning which is experientially based. Shop and laboratory experiences should be brought in to provide a counterbalance to the otherwise heavy verbal learning exposure.

I'd like to enlist your support in this kind of endeavor, and those of you who are going ahead with research that relates to the noncollege bound or the disadvantaged student ought to be involved. We are just beginning to go into the field with requests for proposals to help structure this enterprise, which will be organized, we hope, in a systems-like manner with perhaps even a PERT program that allows us to identify the critical actions which must take place at certain times along the way. We will be looking to the universities for considerable help in field testing and trying out these materials on students. We are hoping to identify several types of schools and school districts, ranging from rural to urban, from advantaged to disadvantaged, to serve as pilot schools. As this unfolds over the next 5 year period, there is going to be a lot of interchange of ideas with people in the field. We hope to keep communication channels open so that we can actively involve those who want to be involved.

Let me just close my remarks by a quotation from Secretary Gardner, who says, and I quote, "No society is likely to renew itself unless its dominant orientation is to the future. A society capable of renewal may not only welcome the future and the changes it may bring but believes that it will have a hand in shaping that future." Thank you very much.

PRESIDENTIAL ADDRESS

Richard M. Clark, State University of New York at Albany

The theme of our Convocation is, "The Emerging Research Worker and His Needs." What "picture in the mind" does this theme create for you? Do you see something akin to the Ajax White Knight riding down the halls of your school on a giant computer and instantly transforming the classrooms with his research findings? Or perhaps you see a kind of squalling infant, being born amid some pain and requiring much care and attention before he is of use to anyone. Does the emerging research worker that you visualize live professionally in an ivory towered university, on the fourth floor of the State Education Building, or in an office in a public school?

As one now working in higher education I am interested in the program of preparation which would be appropriate for an educational research worker to be employed by the public school. Just exactly what should he be prepared to do? I would like to consider with you today five roles that the emerging research worker might play in the public school setting. I do not suggest that these roles are mutually exclusive, but I do believe that each role may require special training experiences.

One role which the emerging researcher might play is that of a "system-free truth seeker." Hilgard, in his essay *A Perspective on the Relationship Between Learning Theory and Educational Practice,* in the 1964 NSSE Yearbook traces six steps on the road from pure science research to established educational practice. His third step, still classed as pure science research, states, "Research on learning that is relevant because the subjects are school-age children and the material learned is school subject matter or skill, though no attention is paid to adapting the learning to school practices." A "system-free truth seeker" might operate at Hilgard's third level and use the school as a kind of laboratory. Pupils, teachers, and administrators would be his experimental subjects as he seeks for truth and knowledge. As Hilgard states, "Of course there may be bridges from any pure-science project to a practical one: perhaps drugs discovered in brain studies of rats may aid remedial reading, studies of interference may suggest intervals between classes or what should be studied concomitantly, and language-vocabulary results in a pure context may guide language acquisition in schools. The main point is that the scientist has not committed himself to relevance."

Is the appropriate educational researcher for the public school setting one who does not commit himself to relevance? Should we prepare a person to use the school as a laboratory in a search for truth and knowledge with the understanding that we expect no immediate utility from anything he does? Is a system-free truth seeker the kind of emerging research worker we should be preparing? I leave the answers to these questions to you, with one related point. Our new campus at the State University at Albany has numerous towers, ivory or otherwise. For those of us who dwell in these and similar towers there is great appeal to the idea of developing system-free truth seekers. We like to impress upon students that theory is important, that profound ideas may take years to become meaningful in terms of applications, and we may even subtly suggest that what is immediately practical is likely to be trivial. Perhaps what we in higher education like happens to coincide with what public schools need, but perhaps it doesn't. An examination of other possible roles may help us to decide to what degree the emerging research worker should be a system-free truth seeker.

A second possible role for the researcher in the public school is that of "educational accountant." At present most schools can tell us more about the fuel oil that goes into the burner in the basement and the efficiency with which it is utilized than they can about the units of arithmetic taught in the third grade and its efficiency for pupils. Perhaps an educational researcher should be a kind of bookkeeper, skilled in maintaining precise records of inputs and outputs for each pupil in the school, and competent in letting us know just how the books balance. Let me attempt to illustrate more precisely what I mean. Suppose we should analyze the records available on a sixth grade boy whom we now call a "retarded reader." We are likely to find an attendance record, a series of achievement test scores, (but usually not the tests themselves) one or two intelligence test results, copies of reports to parents, a capsule comment each year from the teacher, and perhaps the record of an hour or two that he spent with the school psychologist. This child has typically spent more than 7,000 hours of his life in school by the sixth grade. Do we really know enough about what has happened to him? Could we develop a specialist who could provide us information about the specific instructional materials used with the child, the specific types of classroom climates to which the child has been exposed, the kinds of reinforcements the pupil has encountered, the cognitive styles which he has displayed in different learning situations, the mechanisms which he has used to handle frustration, and the ways he has interacted with his peer group? The role that I see for the educational accountant is descriptive. He would provide a detailed picture of what goes into the education of each child, and what behavior comes out. Others in the educational setting would use his data in developing program changes, in guiding the child, and in making various kinds of operational decisions.

Perhaps the role that I have attempted to sketch for the educational accountant does not seem to you to really involve educational research. Please note that the role will demand the development and validation of much instrumentation and the use of many research-related skills. Perhaps the crucial question is whether a person playing the role I have described in the public school setting would be of sufficient value to justify his position. If so, do we know enough to train such a person, and is he the one we will call an "emerging research worker?" Again, let us consider some alternatives.

A third kind of emerging research worker might be what I have chosen to call a "grant-getting go-between." All of us have been exposed to various public and private sources of money available merely for writing the proper proposal. We are also aware that many proposals written during the past year were developed by professionals in the school with no special training and no special interest in developing research studies. Often these proposals involved good ideas and good procedures for carrying out the ideas, but did not provide adequate means of evaluation. Perhaps the emerging research worker should be an expert in knowing where research money is and how to get it. He might be the one who helps the school psychologist develop his proposal for the National Institute of Health, who works out the Title I proposal for elementary school guidance, or who finds foundation support for a new reading program. In each case the "grant-getting go-between" would be working with specialists in the school to put their ideas into a researchable format and finding support for that format. To show that we in professional education really believe in the free enterprise system, we might even pay our "grant-getting go-between" on a commission basis —perhaps three per cent of all grant monies from public sources and double pay for private money.

At the moment at least, money is not the problem. Rather we need more rigorous evaluation of new educational ideas than the criterion, "Parents seem to like it." Perhaps the emerging researcher should be one who inserts a valid research element into rapidly developing new programs so that in time we may really know the worth of these new programs. Is the "grant-getting go-between" the kind of researcher we should train for the public school? Again, let us consider another alternative.

Perhaps our emerging research worker might be a "Foreign Language Interpreter and Translator—FLIT." School people could then revive the old advertising slogan, "Quick Henry! The Flit!" in a new context. To illustrate that, for many potential consumers, educational research reports are a kind of foreign language which requires translation, I selected a study with a simple sounding title which seemed relevant to the concerns of a public school teacher. The study, "Adolescents' Acceptance of Authority," appeared in the *Harvard Educational Review* in the summer issue of 1965. The author states in his conclusion:

> Having conceived of authority as the acceptance of the justification of power relationships, we have found some evidence that, for a certain population of adolescents, such acceptance derives from commitments to distinguishable criteria. We conclude that subjects' verbal acceptance reveals criteria which have been formerly hypothesized by political and sociological theorists, but which have not been investigated in a "clinical" sense. This clinical application indicated that additional criteria might be needed to describe rationales which people use as vehicles to accept decision-making power (in the area of public controversy).
>
> The differential frequency with which the criteria were invoked indicated more or less popular criteria within this adolescent sample. Correlations between the frequencies of the criteria suggested that some of the criteria are meaningfully related to each other in terms of familiar political and psychological observations.

"O.K., so what else is new?" the high school English teacher may ask. (In unguarded moments English teachers can be quite colloquial) "I'm having enough trouble with 'Winston Tastes Good Like Your Cigarette Should' and the symbolism in Macbeth to worry about jargon like, ". . . Correlations between the frequencies of the criteria. . ."

This teacher needs FLIT—a person who can help to bridge the gap between research and practice by translating research findings into a language that can be understood by other professionals in the educational setting. In my opinion, the study which I have quoted is a good one with possible implications for educational practice. However, such implications need to be translated from research findings before they have meaning, and this is the task I see for the Foreign Language Interpreter and Translator. Such a person might help to reduce the often lamented gap between research findings and application.

The last role that I would like to consider for the emerging research worker is that of "public relations expert statistics spouter—PRESS." This person would have at his fingertips facts and figures about the educational enterprise which would be available for all kinds of press releases. He could show the PTA, with colorful graphs, why a new school building is necessary, convince the Board of the need for higher salary schedules, and demonstrate to the local John Birch Society that the citizenship education program is really effective. In this day and age of image building, PRESS could lend an air of scientific respectability to the school.

I have suggested that the emerging research worker might be a system-free truth seeker, an educational accountant, a grant-getting go-between, a foreign language interpreter and translator, or a public relations expert statistics spouter. You may have other roles which you see for the emerging research worker. The important question is to determine what we really want such a person to do in the public school and then decide how we can best prepare him to do it. The "we" I have in mind for this decision-making is a joint "we" including public school personnel, state education department, and higher education. Decisions about function and role need to be made before we decide whether the emerging research worker needs teaching experience, courses in statistics, or work in public relations.

It is my hope that groups such as ERANYS and convocations like this one may bring about a formal and an informal discussion which will bring us closer to needed decisions.

THE EDUCATIONAL RESEARCHER: COMPUTER SCIENTIST OR RENAISSANCE MAN?

John B. Carroll, Harvard University

Among other things, life is full of assumptions, and I would like to begin by announcing one assumption about which I have a certain amount of doubt, but which it is *necessary* to make if I am going to get on with this talk—namely, that there is such a field as educational research and that it has something useful to do other than to improve its own methodology. Several years ago I taught one of my graduate classes in educational research to distinguish among several varieties of research: methodological, measurement, survey, correlational, and experimental treatment research. I then sent them to the library to classify articles in selected educational research journals. Nearly 50 percent of the articles scanned were of a purely methodological character involving no empirical data,—dealing with such things as new statistical formulas, new methods of scoring tests, or new methods of computation and data handling. Perhaps the situation has improved in the last few years, but at the time I was struck with the enormous fondness that educational researchers seemed to have for statistics and methodology; they were almost afraid to touch ground and get their hands dirty with data. Obviously, developing new methodologies and new statistical formulas is not going to change the world very much, nor even influence the process of education, which I take to be the proper concern of educational research.

Nowadays a new image of the ideal educational researcher is emerging in some quarters,—the image of him as a "computer scientist,"—a person who not only knows his educational research methodology and statistics, but who is also skilled in programming computers to handle large quantities of data, to run educational experiments "on line," or to give "computer aid" to instruction. (That phrase, *"computer aided instruction,"* makes me wonder sometimes whether the computer is giving first aid to something that is in emergency need of repair!) Let me hasten to say that many of the new developments in research methodology, statistics, measurements, computer aided instruction, data banking, systems analysis, simulation techniques, and so forth are promising, and may pay off in the long run. I have felt it important to encourage and actually participate in some of these developments. It is clearly necessary for the science of educational research to perfect its methodology and its tools,—both of the soft and the hard varieties. There has been strong reason for the new breed of educational researchers to develop themselves toward the image of the computer scientist I have been depicting.

When I say "computer scientist," I use this phrase only as a shorthand label, as it were, for a wide class of educational research specialists. Not all of those I refer to in this way are necessarily computer scientists in the strict sense; some of them will never get near a computer or even an IBM card—except on receiving a Treasury Department check for reviewing Office of Education proposals. My phrase connotes the hard-nosed methodologists, the statistician, the test maker, the "measurement man," the specialist in educational evaluation, the mathematical modelmaker, the specialist in audiovisual technology or instrumentation, the decision theorist, or the systems analyst. As educational psychology develops, we will undoubtedly see more subspecialties of "computer science" in education emerging. This is all to the good; practitioners of these subspecialties can make important contributions to education.

The development of educational research and technology through computer science and allied methodologies has its distinct limitations, however.

One of these is the rather well known and much discussed fact that there are limits to what can be done with exact and highly controlled experimental design in the schools. A Laboratory of Experimental Design has been established at the School of Education at the University of Wisconsin, but we have yet to see any major educational or instructional question that has been treated successfully by the techniques so well developed by that Laboratory. When one contemplates research on a major question, like how language laboratories should be used in foreign language instruction, or under what conditions homogeneous grouping is superior to heterogeneous grouping, or in what way class size is a critical factor in instruction, he is faced with so many difficulties and necessary compromises in applying classical Fisherian experimental design that the effort seems hardly worthwhile. But this very limitation of highly controlled experimental design leads to a paradox: the person who tries any kind of educational research at all,—even though unavoidable circumstances may prevent it from going by the book, —needs to be acutely aware of the technical subtleties of experimental design and the possible threats to internal and external validity so that he at least knows what controls he is failing to employ and what ambiguities in interpretation lurk in his research plans.

But highly sophisticated research employing exact experimental design and carefully validated measurements has an even graver limitation, and this one is inherent in it,—namely, the fact that even the best experimental design is only a tool for testing good ideas; it cannot create ideas of any sort, except insofar as its *results* may suggest good ideas to the mind of the researcher. The good ideas must come from somewhere. Often they come from educational practitioners who have no idea of educational research design. Sometimes they come from theoreticians who may or may not be educational research specialists. But they rarely come from people who are purely methodologists or "computer scientists." The educational researcher who conforms too closely to the ideal of the "computer scientist" runs the risk of being written off as a pure technician, somewhat analogous to the "paramedical specialist" who is ancillary to the physician.

Furthermore, even when educational research of a sophisticated character is testing a *great* idea, the dull statistical tables that result have little public relations value or power to move men. Jerome Bruner's purple passages in *The Process of Education,* bereft of any suggestion of statistical technology, have had much more influence on the course of education than his more detailed and carefully designed studies of thinking processes. Even the recent study of differential "educational opportunity," ordered by Act of Congress and conducted with enormous sophistication in sampling design and educational measurement technology by a group under the direction of James Coleman, will probably not have its full impact until its findings can be translated into a book with the emotional power of, say, Rachel Carson's *Silent Spring.* (This would be a wonderful project for some gifted writer, wouldn't it?) Under the best of conditions, scientific educational research is difficult to convert into statements that can be useful to the educational practitioner and that will also move him to action.

These and other considerations yet unmentioned lead me to suppose that the educational researcher in the guise of a computer scientist is likely to be a rather ineffective and, in the long run, frustrated character. He will not have the impact on education that he or we may desire. How can he improve his position? My answer, as you may have already divined, is implicit in the title of this talk: he must become a Renaissance Man. Before you conclude that all educational researchers must become Leonardo da Vincis, let me say that again I am using a phrase not as an exact description, but as a symbol for a concept. The Renaissance Man is not only a technician and a methodogist; he is also a broadly trained scientist;—and not only that,—he is a humanist who appreciates and understands, to a considerable depth, the whole spectrum of human knowledge. In short, he has a sound liberal education in addition to his technical training. A modern Leonardo da Vinci might be the model or the ideal for our Renaissance Man *qua* educational researcher, but we cannot expect such versatility and competence in very many. Nevertheless, the requirements I have in mind are high and demanding. Consider some details of these requirements.

First, the educational researcher who approaches the ideal must be a practical man of sorts. He must have a familiarity with schools and their problems that he cannot acquire merely by a few desultory visits. He must be able to see the problems of the school at least partly from the standpoint of the teacher or the educational administrator. He must be able to formulate educational research that will be realistic in terms of the practicalities of running schools. He must understand and get along well with school people in order to gain entrée for his surveys and experiments. He must have some knowledge of the politics of education and the strategies and tactics of introducing educational improvements. (I was about to use the fashionable term "innovations," but I would be prejudiced against an innovation unless it is indeed an improvement: and the sad fact is that there are many innovations that are not improvements, and many possible improvements that are not innovations, but have been waiting around for decades.)

Secondly, the educational researcher who approaches my ideal of the Renaissance Man must have an intelligent understanding of the whole realm of curriculum as taught in the school. If he is going to do research on the teaching of the "New Math" or the "New Science" or the "New English" (and I don't mean the New England dialect!), he has got to understand the goals of these new curricula and be able to communicate with their purveyors without making himself appear silly or otherwise unacceptable to them. This requires a considerable degree of mental flexibility and willingness to give up old modes of thinking. It is hard enough for teachers of mathematics, science, or English to learn new ways of thinking about their subject matter; it is even harder for the educational researcher, who approaches these subject matters from the outside, so to speak, to do this. It is all too easy for the test maker or curriculum researcher stubbornly to insist upon the sublime propriety of what he was taught or the way he was taught, 20 or 30 years before, rejecting the new wisdom of the curriculum developer. For example, an educational researcher who had experienced only the traditional methods of teaching English grammar would fail to comprehend contemporary thinking in this field and would probably be unable to perform meaningful research in it, even though otherwise he might eventually be able, let us suppose, to show a superiority for the traditional method over methods being promoted currently.

Thirdly, the educational researcher nowadays must be not only a methodologist, but also to a large extent a behavioral scientist. Because his research deals with changing human behavior, he has got to have a deep knowledge of theory and research in those aspects of behavioral science that can help him in this task. Sociology, anthropology, political science, economics, and even history may be useful to him, but above all, I think, psychology,—because the educational process is concerned centrally with the individual rather than the group. And within psychology, the most useful knowledge probably comes from research on learning, motivation, attention, maturation, and other basic topics. It is not without reason that a large number of educational researchers have a primary affiliation with psychology. I do not insist that the educational researcher be a productive scholar in a behavioral science; his use of behavioral science is more generally in application than in discovery. Nevertheless, many of the most distinguished educational researchers, like E. L. Thorndike, have made genuinely creative contributions to a behavioral science. Because behavior theory still seems so inadequate in its application to education, and because the prime application of behavior theory would seem logically to fall within education, it behooves the educational researcher to think twice before deciding not to engage, at least to some extent, in fundamental research in behavioral science.

Fourthly, the educational researcher who emulates the Renaissance Man must be a creator or an inventor in a sense beyond mere productive scholarship. He must create new educational ideas, to be tested by his methodology, and guided by his research. He must invent new educational materials and techniques to be tried out in the schools and colleges. With his knowledge of the practical problems of education, of curricular trends, and of the principal theories and results of behavioral science, he has at least as good an opportunity as anyone else to initiate new and useful educational proposals. If he is also an essentially creative person, whatever that may mean or entail, education will be that much the gainer.

And fifth, I should like my Renaissance man-educational researcher to have a flair for communicating his ideas and findings in speech and writing in such a way that they can gain the wide audience that they need to gain if educational research is to have its rightful place in the scheme of educational progress. It is all very well for him to pile up an impressive number of research publications reporting empirical findings in the manner approved for the *Journal of Educational Psychology* or the *American Educational Research Journal,* but there is also a place for the thoughtful, inspiring public address or printed essay that will reach the minds of teachers, curriculum supervisors, school administrators, or the lay public. The requirement is not that he be a literary genius; he needs only to be capable of clear thought expressed in an interesting and imaginative expository style.

These, then, are my modest minimum requirements for a Renaissance type of educational researcher. I remind you that earlier I said he must also have a strong streak of the computer scientist in him. Beyond this he is a free man! He may make his own choices about cars, clothes, beards, marriage, politics, religion,—and the ways he would spend his leisure time if he had any!

You may say the ideal I've depicted is an impossible combination of almost incompatible skills, knowledges, and talents, or that it would take a lifetime of training to produce such a person. Renaissance Men, you may say, were possible in the Renaissance, but not today. I will grant that the standard is high and that the educational researcher who approaches it will be, at least for the time being, a rarity. Further, I grant that the time required to produce a really competent educational researcher is long, but probably it is no longer than the time required to produce real competence in any other professional field. This is simply something that we must reckon with and plan for, better than we have in the past. I would like to state some of my thoughts on the selection and training of educational researchers.

Obviously, a central requirement is high intelligence and high aspiration. Among the millions of undegraduates in our colleges today there must surely be a sizable number of topnotch students who might be attracted to educational research rather than to nuclear physics, international politics, or what have you, provided the potential rewards and satisfactions of excellence in creative educational research were made known to them. A surprising number of outstanding educational research workers of my acquaintance today had undergraduate majors in physical or biological sciences. This is borne out by Bargar's survey of educational researchers, which found that 14.2 percent of 3737 respondents had undergraduate majors in these two sciences.[1] But Bargar also found that 37.8 percent of his respondents had undergraduate majors in psychology, sociology, or other social sciences, and, I suppose, these are the fields where we would be most likely to find candidates for advanced training in educational research. What we would have to be most careful about, however, is the danger that out of all the potential candidates, educational research would attract only the less able,—the ones who might think of education and educational research as a "soft" kind of work, not as challenging as pure academic research in a discipline like mathematics, psychology or chemistry.

But while we are considering the undergraduate phases of the development of an educational researcher, let me plead for breadth of study,—that is, for a liberal education in all branches of knowledge. It is in this period, supplementing a good secondary education, that the prospective educational researcher can acquire knowledge of subject matters that he will need when, later, he undertakes research in the teaching of one or other of these disciplines. The study of mathematics, statistics, and science are obviously basic to his later involvement in research methodology. Probably he should acquire a considerable competence in one or two foreign languages, chosen among French, German, and Russian, because of the increasing importance of educational research publications in those languages and the increasingly international look of educational research. And he would need to be well grounded in literature, art, and philosophy, because the humanistic areas are so important in the school curriculum. A major or a minor in a behavioral science would be a desideratum but not a necessity. I see little reason, however, for him to study education at the undergraduate level, except possibly one course in educational psychology or the history of education, taken as an elective.

An excellent plan for the first year of graduate study might be one spent in preparation to be a teacher,—as in one of the many MA in Teaching programs that now dot the academic countryside, followed by a year of actual teaching. Alternatively, the student might spend two years in some sort of work such as teaching in the Peace Corps or the Job Corps. The period immediately after the undergraduate years would be a good time to acquire teaching experience unencumbered by the doubts and qualms that would be engendered by later graduate study, and my prescription for the Renaissance Man-educational researcher certainly requires some teaching experience. But there are two dangers in planning teaching experience at an early stage after the bachelor's degree: one, the student may lose continuity between his undergraduate study and his later, advanced graduate studies, and two, there will be a certain number of people who may not stand up to the rigors of early teaching experience even though they could later qualify as outstanding educational researchers. It would be a pity to lose these people on their way to the educational research profession.

On his return to graduate study after a stint of practical teaching experience, the prospective educational researcher must, I am afraid, be prepared for a fairly lengthy and arduous course lasting 3 or 4 years before obtaining the doctorate. Normally the training will be in one of the behavioral sciences, but with a considerable infusion of courses in education: philosophy of education, curriculum development, educational administration, counseling and guidance, teacher supervision, principles of teaching, and the teaching of selected subjects in the curriculum. If the specialization is in educational psychology, there must be a heavy concentration of courses and readings in various aspects of that subject, and of course, in any case, there must be a firm mastery of some of the more advanced and esoteric aspects of educational research methodology, including experimental design, multivariate statistics, theory of measurement, and the like. The "computer scientist" streak of the Renaissance Man must be nourished with courses and practical experience in computer programming and computer aided instruction.

---

[1] Bargar, Robert R. Who is the educational researcher? Pp. 13-47 *in* Guba, Egon, and Elam. Stanley (Eds.), *The training and nurture of educational researchers.* Bloomington, Indiana: Phi Delta Kappa, 1965.

Towards the end of the trainings, the student must be introduced to actual educational research work on an apprenticeship basis, working as a project assistant. But he should be encouraged to strike out on his own, still under supervision, of course, to the point of developing research material publishable under his own name.

The course of training should include lectures, readings, and practical work on the organization and administration of educational research, including the supervision of research staffs, the preparation of project proposals and budgets, and "grantsmanship" in general. Similarly, there should be supervised work on curriculum development and evaluation or some other grassroots kind of problem.

I would insist that there be a dissertation that would constitute an original piece of educational research, either as an independent project or as part of a larger project. It would be a contradiction for an educational researcher to obtain his doctor's degree without performing educational research.

To round out the training of the educational researcher, I would strongly urge that post-doctoral research experience, again under supervision, be a normal and expected part of the researcher's career. Usually he would take his post-doctoral training at a university or research bureau other than the one from which he had received his doctorate, in order that a different staff could inject him with fresh ideas.

Such a background would go far, I think, to produce the Renaissance Man that is my ideal. It is only a logical extension of some programs that are already in operation. At the same time, I have emphasized throughout that the standards must be set high and that the training period must be lengthy. My aim has been to show that actually the requirements for the training of the effective educational researcher are perhaps higher and more demanding than they are for most other types of specialists. I think, for example, that they are higher than for the normal Ph.D. in psychology. Failure to recognize this fact has plagued the field of educational research because it has resulted in inadequate planning for pre-doctoral and post-doctoral fellowships, research assistantships, and the like. As another consequence, we have turned out too many research personnel who are very far from being Renaissance Men;—either they have been strictly computer scientists and methodologists, or they have been insufficiently trained and ineffectively operating generalists. This is responsible for the poor and unexciting quality of a great deal of educational research: poorly conceived, poorly designed, poorly performed, and poorly reported. It is no wonder that educational research has been in bad repute.

Let us, therefore, try to train more Renaissance Men in educational research.

Highlights of Recent Research

SELECTED STUDIES IN ADMINISTRATION: RECENT RESEARCH IN THREE SOCIAL SCIENCE DIMENSIONS

Laurence Iannaccone, Frank W. Lutz, & Harry J. Hartley, New York University

This paper has two main purposes. One is to identify some of the theory based research in educational administration, recently done, now underway, or predictably to be undertaken soon. Its second purpose is, chiefly by way of illustration, to depict the process by which theory based research contributions come to be made and eventually have an impact on the training of educational administrators. Research in selected aspects of three social sciences applied to the study of educational administration will provide the vehicle for illustration.

A process which we shall inaccurately dignify beyond its merits as philosophizing about specific problem areas in educational administration seems to take place first. Then some loose historical studies, some more or less—usually less—theory based descriptive work, and much more census taking called *research* takes place. If the problem area continues to evoke interest, conferences of professors frequently talking about the need for theory based research and the potential contribution of the social sciences may overlap and follow the census taking and descriptive work. Sometimes worthwhile theorizing goes on at a few of these. In the meantime, bright graduate students not invited to the conferences may, one here, one there, begin to produce theory based research in the area. If these dissertations begin to cluster around a few scholars, as in even a single college department of administration, the move to somewhat larger theory grounded research programs may begin. This may take place if one or more of the graduate students doing a theory based dissertation becomes a research professor and with his student follows up his dissertation. At such places the introduction of new courses and program revision in educational administration may begin. At any one of these points or stages the development may, however, stop due to a wide variety of factors including chance areas.

The area of research on teacher organization and negotiation has been selected to illustrate the first phase of this process as it is now being approached with sociological theory which is beginning to produce theory based research.

The Sociological Dimension, Research On Teacher Unions and Associations

The field of teacher negotiations is too new to have produced much research. Unfortunately, work done in the public and private sectors is not directly applicable to teacher groups. Considerable attention has been given to the area of public employees with emphasis on civil service, Federal, State and local. A recent book by Walton and McKersie[1] has developed an extensive theoretical analysis of the negotiation process. This work is generally pointed at the private sector. Its theoretical base rests largely in psychology, and little research in the actual operation of unions is incorporated.

Because of recent problems in educational administration, there has developed a considerable writing on the subject. Some of the writers are not new to the area. Myron Lieberman[2] produced his first book in 1956. His work has contributed to what might be called the "philosophical discourse" but has provided little generalizable knowledge about teacher militancy. In general, most of the work published in the journals of the profession has furthered this philosophic dialogue or described the present status of the situation. In addition, journal articles and National Education Association "research" have contributed to the "social census taking" in the area. Thus, data are available that give the number of associations and union members, the number of written contracts and agreements, and the kinds and extent of sanctions and strikes, etc. Two recent books, one from the association point of view,[3] the other somewhat broader,[4] describe in an atheoretical way the present status of the field. Another book that calls upon the collective knowledge of six persons in different disciplines attempts to develop theory that might explain the behavior of teachers and teacher groups in the present administrative and organizational setting.[5] While this last volume provides an attempt to utilize theory in understanding the problem of teacher power, little research is cited except as related to state politics.

---

[1] Richard E. Walton and Robert B. McKersie, *A Behavioral Theory of Labor Negotiations,* (New York: McGraw-Hill Book Co., 1965.)

[2] Myron Lieberman, *Education as a Profession,* (Englewood Cliffs, New Jersey: Prentice-Hall, 1956.)

[3] T. M. Stennett, Jack H. Kleinmann and Martha L. Ware, *Professional Negotiations in Education,* (New York: The Macmillan Co. Limited, 1966.)

[4] Myron Lieberman and Michael H. Moskow, *Collective Negotiations for Teachers,* (Chicago: Rand McNally and Co., 1966.)

Perhaps the most systematic research done in this area has been done by Wesley A. Wildman and his associates at the University of Chicago.[6] Again, this cannot be called theoretical research. It focuses heavily on the legal aspects of the situation and on a descriptive presentation of what exists.

Conferences across the country have also used this area as a point of emphasis. Excluding the local, and county meetings, one can point to the conferences held at the Universities of Rhode Island, Indiana, Arkansas, Pennsylvania, Oklahoma, and New York University. These last two meetings resulted in publications that include some theoretical approaches, but little research.[7] Undoubtedly these efforts at theorizing and the recent work of sociologists in industry such as Blau, Etzioni, Gouldner, etc., will result in theory oriented study in this area in the near future. As for now, the shelf is nearly bare.

Presently most of the theoretically based research concerned with teacher unions or associations has been done by doctoral candidates with the result that once done, little dissimination occurs. Ralph Schmid[8] studied the historical aspects of the NEA as the AASA, the teachers union, and the classroom teachers' association struggled for the reins of leadership. Steven Rubin[9] interested in the effect of spacial change on the sentiments of individuals within the informal structure of the elementary school, studied the sentiment change toward a teacher union. Seymour Evans[10] studied the organization of schools in relation to bureaucratic theory and found that the predisposition of teachers, along a Gemeinschoft-Gesellschoft continuum, was related to their tendency to seek organization along a bureaucratic-prebureaucratic continuum. This he notes is related to the tendency toward union militancy or loose association. Other doctoral research could be cited. These three only serve as indicators of the possibility of theory related research in this area.

Underway at the present time are some interesting research projects which indicate an attempt to study this area of educational administration. A project funded by the Center for Urban Education is looking at the possibility that Gouldner's bureaucratic sanction model can account for principal leadership behavior in schools operating under union contracts. Another study, an outgrowth of Evans' work, is attempting to test the hypothesis that teacher membership is related to Gemeinschoft-Gesellschoft predispositions and organizational pressure. Two other studies are looking at the role of the superintendent in negotiations: one uses role as a dependent variable related to the independent variables of district conflict and size; the other uses role, along with district size and fiscal structure, as independent variables related to board, superintendent, and teacher satisfactions with the negotiation process in the district. All of these projects are related to a central research theme at New York University and should result in integrated systematic theory building in an area of education.

Finally, it may be noted that this summer a workshop will be taught by Evans on the principal's role.

Attention will next be given to the political dimension focusing on recent research in the state politics of education and emphasizing the later stages of the process, the theory based research program, and its diffusion through program changes in the training of school administrators.

The Political Dimension, The State Politics of Education

Following some of the early phase philosophizing about politics—"t'is—"t'ain't" in education, a number of atheoretical descriptive studies were produced on the politics of education at the state level.[11] As indicated at the outset of this paper, these need not necessarily have resulted in theory and theory based research. They were, however, not only widely read, but also studied and at one institution, at least, they functioned as the basis for a four semester sequence of research seminars in the politics of education. This sequence in turn produced theory in the state legislative politics of education,[12] publication of this and theory based research in the form of dissertations on the state of California's politics of education.

---

[5] Frank W. Lutz and Joseph J. Azzarelli, *The Struggle for Power in Education* (New York: Center for Applied Research in Education, Inc., 1966.)

[6] Wesley A. Wildman, "Legal Aspects of Teacher Collective Action," *Theory into Practice,* IV (April, 1965, 55-56.)

[7] Robert E. Ohm and Oliver D. Johns, *Negotiations in the Schools* (Oklahoma City: University of Oklahoma, 1965.)

[8] Ralph D. Schmid, *A Study of the Organizational Structure of the National Education Association 1884-1920,* (Unpublished Ed.D. dissertation, St. Louis, Washington University, 1963.)

[9] Steven Rubin, *A Social System Analysis of the Informal Structure of a School Faculty: An Investigation of the Process of Sentiment Change* (Unpublished Ph.D. dissertation, New York: New York University, 1964.)

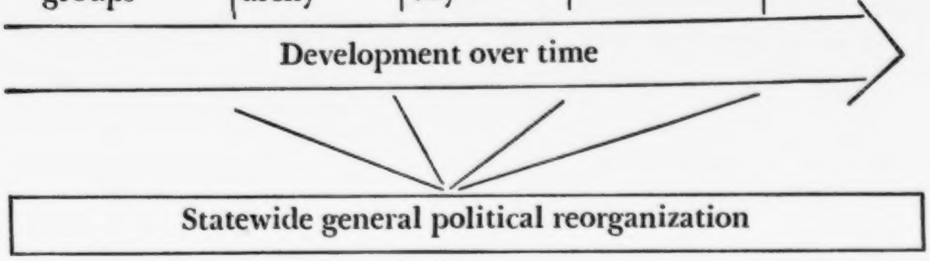
[10] Seymour Evans, *Towards A Theory of Teacher Collective Organizational Behavior* (Unpublished Ed.D. dissertation, New York: New York University, 1966.)

[11] Stephen Bailey and others, *Schoolmen in Politics,* (New York Institute of Administrative Research, Teachers College, 1963) Michael Usdan, *The Political Power of Education in New York State* (Syracuse: Syracuse University Press: 1962.)
Nicholas A. Masters, *State Politics and the Public Schools,* (New York: Alfred A. Knopf, 1964.)

The earlier descriptive studies were treated by *post factum* analyses. Out of this emerged a typology based upon characteristics of the organizational structure displayed at the key linkage point and a model of the prevailing pattern of state educational politics.[13] Subsequently, the taxonomy was refined into a developmental construct. A diagrammatic representation of the developmental construct follows:

A Developmental Construct of The State Politics of Education

	Phase one	Phase two	Phase three	Phase four
Characteristic 1. Structure	locally based disparate	Statewide monolithic	Statewide Fragmented	Statewide syndical
2. Political behavioral	entrepreneurial	cooptational	Competitive	coalitional
3. Leadership groups	squirarchy	oligarchy	polyarchy	synarchy



This became the basis of a series of longitudinal studies on the State politics of education in California.[14] These, in turn, partly tested the utility of the construct and elaborated it. Finally, courses in the politics of education influenced by this work exist at the Claremont Graduate School, Auburn University, the University of Wisconsin and New York University.

A similar process may be seen at the local district level with respect to the politics of education. Descriptive work by Lutz,[15] followed by theoretical elaboration using other works in local government led to theory related to political change resulting in specifiable patterns of superintendent turnover and succession.[16] These verificational studies involving in all 161 school districts, provided tests of hypotheses derived from the theory. In turn, this work now influences the courses cited earlier and hence, the training of future school administrators.

---

[12] Laurence Iannaccone, *The Future of State Politics of Education*, Frank W. Lutz and Joseph J. Azzarelli eds. in The Struggle for Power in Education (New York: The Center for Applied Research in Education, Inc. 1966.)

[13] Laurence Iannaccone, *The State Politics of Education* in a publication of the Center for Research and Leadership Development in Vocational and Technical Education, (The Ohio State University, forthcoming.)

[14] Among which see, B. Dean Bowles, *Conflict and Consensus in the State Politics of Education;* Donald N. McIsaac, *A Statistical Analysis of the State Legislature in California;* W. Deane Wiley, *Political Interaction of Education and the California Legislature: 1849-1963;* (Unpublished Ph.D. dissertations, The Claremont Graduate School, 1966.)

[15] Frank W. Lutz, *Social Systems and School Districts, a Study of Sentiment and Interaction of a School Board,* (Unpublished Ed.D. dissertation, Washington University, 1962.)

[16] Robert M. Freeborn, *School Board Changes and Involuntary Turnover;* Richard S. Kirkendall, *Discriminating Social, Economic and Political Characteristics of Changing vs. Stable Policy-Making Systems in School Districts;* Eldon G. Shafer, *Unification: A Change of Power Structure Reflected in Board Composition and Superintendent Selection;* John C. Walden, *School Board Changes and Involuntary Superintendent Turnover.* (Unpublished Ph.D. dissertations, The Claremont Graduate School, 1966.)

The next dimension, the economic, raises an interesting question. Of the three dimensions, it should logically have had the greatest impact already, because it is a social science more detailed, specific, and theoretical than the other two and yet its potential contributions to educational administration have yet to be well exploited. But a beginning has been made here too.

The Economic Dimension

A school may be conceived as an economic system, as well as a social or political system, for we are merely examining the same structure with different analytical tools. Recent published scholarship in this area of educational economics indicates that this discipline can contribute to theoretical construct development and research in school administration. Of all the social science disciplines, economics appears to possess the most highly developed theoretical structure. Economic rationality incorporated mathematics and systems analysis techniques as a means for relating measurements research to decision making. The discussion which follows contains a brief description of three economic concepts which may be applied to administrative research: systems analysis, program budgeting and economic value dimensions.

Systems Analysis

A topic of central concern to both economists and administrators is optimum allocation of scarce resources. Economics provides consistent, inclusive theories of decision making under which allocation problems can be restated in system analysis terms and techniques. The purpose of systems analysis is to improve methods of decision making. Maccia[17] developed guidelines for research involving this concept in education, although the Rand Corporation provided earlier data. In reviewing the research literature on systems analysis, one may encounter similar, yet technically different, terms such as operations research, input-output analysis, management information system, program evaluation, cost-benefits analysis, and others. Prest and Turvey compiled an extensive bibliography of current research,[18] particularly involving cost-benefits analysis. Thomas is currently conducting research in probability models and systems analysis design for education at the University of Chicago.[19] The implications of a systems approach are great, and research interests in this area are only beginning to expand.

Program Budgeting

The major characteristic of a typical school budget is its incremental nature. This year's budget, for example, possesses the same structure as last year's, except that it may contain five percent increases in the various line-items. Little attempt is made to evaluate the various programs which are supported by the budget. A budget is almost never reviewed as a whole each year in the sense of reconsidering the value of all existing programs. On the other hand, the program budget, which has been successfully applied to the Federal Government by Defense Secretary McNamara, is organized in terms of program elements that produce designated end-products, rather than a line-item, "object" basis. Costs are continually analyzed in relation to the measurable benefits produced by the program elements. Certain municipal governments such as Los Angeles and New York's Nassau County and major school districts such as Chicago, Memphis, and Dade County, Florida are conducting research to determine how to implement the program budget. In a major 1966 study of more than 100 large city school systems, James of Stanford concluded, "It is likely that the use of program budgeting will slowly increase in large city school systems during the coming decade."[20] James also noted that Title I of the Elementary and Secondary Education Act of 1965 places specific evaluation requirements upon local districts which involve cost-benefits analyses and program budget-related information.

Economic Value in Education

This rather global topic includes diverse types of research of which three contrasting examples are mentioned· productivity, individual value preferences, and human capital.

1) *Productivity:* Education is a major determinant of productivity increases and economic growth in this country. Productivity indexes, which consist of a ratio of the quantity of goods produced to the input of costs, have been developed for various industries, but there

---

[17] George S. Maccia, *An Educational Theory Model: General Systems Theory*, Occasional Paper 62-126 (Columbus: Bureau of Educational Research and Service, Ohio State University, 1962.)

[18] A. R. Prest and R. Turvey, "Cost-Benefits Analysis: A Survey," *The Economic Journal*, vol. LXXV, December, 1965, pp. 683-731.

[19] J. Alan Thomas, *Probability Models and Computer Simulation and Preparation Programs for Educational Leaders*, Paper prepared for U.C.E.A. Conference at Columbus, Ohio, May 18, 1965.

[20] H. Thomas James, and others, *Determinants of Educational Expenditures in Large Cities of the U.S.* (Stanford University, School of Education, U.S.O.E. Cooperative Research Project No. 2389, 1966), p. 89.

exists no such index in education at the present time. Benson focused upon studies of educational productivity measures.[21] Beginning with the premise that education is handicapped by not being able to relate changes in dollar value of inputs to some series of outputs, he concluded that some forms of partial measurement may be suitable. Thus, not all of the output of a school must be stated in quantitative terms. As indexes of educational productivity are refined, more precise justification for increased expenditures on education can be given. For example, increased spending could be defended in terms of the higher yield of products that the dollars provide.

2) *Individual Value Preferences:* Economic analysis can be of help to educational leaders who are searching for explanations which underlie the behavior of those who provide the financial support for public education. Economics treats the motivating factors of human activity as economic motives, and provides explanations for the phenomenon of human behavior. One indication of value preferences is the extent to which taxpayers vote for or against referendums for school budgets, bond issues, building approval, and acquisition of sites. In an exploratory study, Hartley analyzed ten value dimensions which eventually may be incorporated into a general economic theory of educational value.[22] One possible implication is that patterns of taxpayer behavior may be modified in the direction of a "core" system of positive educational values after researchers are better able to assess factors which influence an individual's propensity to consume public education.

3) *Human Capital:* Economists and educational finance researchers are directing greater attention to the extent to which man invests in himself. We know that a society can make investments in the productive capacity of men as well as machines, but until recently, the very thought of investment in human beings as capital goods was rather offensive to certain people. Schultz conducted pioneer work which has reached a stage in which a quantitative basis has been laid for the validation of the view that the principal cause of American productivity and growth lies in the large and continuous rate of investment in education.[23] In this "human capital" approach, human beings are conceived as capital assets, or wealth, or resources. As an asset, a person is able to invest in himself and to obtain a measurable yield. A most important form of human investment is education, and the yield, or return, on educational investment is quite high. A part of this yield cannot be expressed in dollar terms, but for the part that can be quantified, mathematical models are being developed. As a result, we know that human capital has been increasing at a rate substantially greater than reproducible (nonhuman) capital. One implication of this research is that it provides economic justification for additional governmental support for education.

Conclusion: Economics has much to contribute to education. It is a science which studies human behavior as a relationship between ends and scarce means which have alternative uses. If the ends are represented by educational excellence generally, and administrative excellence, specifically, then the application of economic theory becomes essential. The several concepts discussed above provide at least partial evidence of the extent to which economic research can be applied to educational administration.

This brief review of three dimensions contributing in a small but significant way to research and training in educational administration suggest two things to us. The use of the social sciences in the study of educational administration can no longer be considered a passing phase. The future contributions of these three dimensions: sociological, political, and economic will be much greater than at present, perhaps even greater than the present authors know.

---

[21] Charles S. Benson, *The School and the Economic System.* Chicago: Science Research Associate, Inc., 1966.

[22] Harry J. Hartley, "Towards a General Economic Theory of Educational Value," *Educational Administration Quarterly,* Vol. II, No. 2, Spring, 1966.

[23] Theodore W. Schultz, *The Economic Value of Education.* (New York: Columbia University Press, 1963.)

HIGHLIGHTS OF RECENT RESEARCH IN BAYESIAN STATISTICS

Donald L. Meyer, Syracuse University

I am restricting this paper to the field of Bayesian statistics and indicating some of the interesting problems which have been, and are currently being researched in this field.

For those of you who are unfamiliar with the Bayesian model, let me do a minute or so of teaching to acquaint you with the method. Bayes' theorem states that the probability density of a parameter θ conditional on data X is proportional to the unconditional distribution or θ mutilplied by the likelihood of the data conditional on θ

or: $f(\theta|x) \propto f_1(\theta) f_2(x|\theta)$

There is no disagreement as to the validity of the theorem. What is argued is its application. In fact, the way in which I stated the theorem might be objectionable to some people because I am implying that a parameter is a random variable, i.e., that it has a probability distribution. For most problems in education, the statistical theory many of us have studied (Neyman-Pearson or Fisher) treats parameters as fixed constants. The only random variables are those which are functions of the sample.

The Bayesian really wouldn't disagree with this except that he finds it reasonable and useful to think of his degree of belief regarding unknown parameters. The $f(\theta)$ in the theorem represents the distribution of belief about sets of values of the unknown parameter. This thinking is not too much different from the way in which people interpret confidence intervals. When we say we are .95 confident that the limits, L_1 and L_2, include the true value of the parameter, it is difficult to think of any useful interpretation other than that based on degree of belief or credibility. Recall, however, that in the Neyman-Pearson theory, there is no foundation for such an interpretation.

A Bayesian analysis, then, begins with the prior distribution of θ and ends with a description of the posterior distribution. Everything relevant about θ is contained in the posterior distribution. Various summary measures might be calculated and reported such as the mean—used as a point estimate—or intervals—used as interval estimates.

Much of the past and current research is concerned with the derivation of posterior distributions for various prior distributions and various sampling designs. The exciting result is that the classical confidence intervals, tests, etc., are derivable from a fairly restricted class of prior distributions. This class is such that the prior distribution for the parameter or a function of the parameter is a uniform distribution. A uniform prior is an "indifference" distribution in the sense that it implies one wouldn't offer odds different from 1 to 1 for any comparison of two sets of values for the parameter. The implication of this is that so-called classical procedures are but a subcase of the Bayesian model. Looked at in this light, Bayesian statistics loses some of its strangeness. Secondly, it means that classical procedures apply when the prior information is essentially zero. One can be a Bayesian and do classical analysis when one finds himself in the "no information" state. More important, however, it forces us to consider the possibility that we do know something about the process under study and therefore, to consider the use of some other prior distribution. As I. J. Good said; "Consider any statistical technique with which you have some sympathy. Find out whether it is equivalent to the use of an initial probability distribution by using Bayes' theorem in reverse. Then replace this initial distribution by a better one."

One of the results in this area concerns the distribution of the mean vector μ, in a multivariate normal distribution. Geisser and Cornfield (1963) used a family of indifference priors for the mean vector which depended on a positive integer, V. The posterior distribution of μ is then one of a class depending on V. If $V = 2$, the posterior probability intervals are the fiducial intervals which follow from Fisher's theory of fiducial inference. If $V = p + 1$, then the posterior intervals for μ are the same as the confidence intervals based on Hotelling's T^2 distribution. One could interpret this result to mean, at least for this situation, that the difference between these two major modes of inference reduces to a selection of a number for V. Furthermore, the marginal distribution of any single mean is $t(N + 1 - V)$. Only if $V = 2$, are the posterior intervals the same as classical confidence intervals. This means that one can choose a prior to give correspondence with the joint interval of Hotelling or one can choose a prior to give correspondence with confidence intervals on single means—but, one cannot do both simultaneously! For those of us who are confused about the relation of multivariate and univariate tests, this fact should be helpful.

Unfortunately, the derivation of posterior distributions from arbitrary priors is mathematically intractable. One of our students, Jim Powers, is working on general computer programs which will allow general priors to be specified for the one-way anova situation and which will calculate the resulting posterior distributions after observation of sample data. Once these programs are obtained, research on the relation between various classes of priors and posteriors can proceed. We will also be able to assess the effect of sample size on the change from prior to posterior distributions.

The Bayesian model is a convenient way of incorporating all information about a process. To give examples, one usually knows something about the measuring instrument used—its reliability, item characteristics, etc. We are beginning to work in this area to attempt to tie experimental design and measurement theory together using the Bayesian approach. Secondly, the problem of cross-validation is one where the initial sample implies a distribution of belief about expected results in the cross-validation sample.

Another example of this type was given by Box and Draper (1965). Suppose there are m fundamental parameters. Observations are taken on a response variable which is a function of some set of the m parameters. Using a different design, observations are taken on some other response variable which is a function of some set of the same m parameters. The problem is to use the information from both designs to make inferences about the m fundamental parameters.

Starting with an indifference prior on the m parameters, three posterior distributions are derived. The first two use, respectively, only the information gained from the separate designs, while the third is the final posterior distribution resulting from consideration of the observations from both designs. By comparing these posterior distributions one can see the relative contribution of each set of observations. One application of this model to education and psychology is to think of the m parameters as fundamental ability constants and the designs as representing different tests and tasks.

Another highlight of recent research is that so-called classical statisticians *are* making use of Bayes' theorem Even though their approach is conceptually different, a Bayesian can feel quite comfortable with their methods and results. One development is called the "empirical Bayes' approach." If we look at the denominator of Bayes' theorem:

$$\int f_1(\theta) f_2(x|\theta)\, d\theta$$

we see that it is the unconditional distribution of X, $g(X)$. This has also been called the predictive distribution of X. If we successively sampled from a process with parameter, θ varying according to $f_1(\theta)$ then we could build up an empirical distribution of X which will approximate $g(X)$. Since $f_2(x|\theta)$ is known, $f_1(\theta)$ is then known implicitly.

An application of this was given by F. Lord (1964) writing in the context of measurement theory. The unknown $f_1(\theta)$ is a distribution of true scores while the empirical distribution, $\hat{g}(x)$, is an observed distribution of test scores.

HIGHLIGHTS OF RECENT RESEARCH CURRICULUM AND TEACHING

Glenn E. Reeling, Jersey City State College

Many innovative aspects have been associated with teaching and curriculum during the past few years. The purpose of this paper will be to describe briefly research which is currently being conducted in two subject areas at the elementary level in the Montclair, New Jersey, Public Schools as well as some of the research which is being conducted in the East Orange (New Jersey) Public Schools.

The first study concerns the Initial Teaching Alphabet (i/t/a) as a method of teaching reading to first grade pupils in the Montclair Public Schools. In this experimentally designed project, 550 first grade pupils in the Montclair Public Schools. In this experimentally designed project, 550 first grade pupils were placed into either an experimental group (teaching via i/t/a) or into a control group which was taught by the traditional orthogonal (t.o.) manner. These groups were matched for the factors of sex and for the scores which were obtained on the *Lee-Clark Reading Readiness Test* in their kindergarten year.

During March of the first grade year, before any of the i/t/a pupils had the opportunity to make the transition to t.o., both groups of pupils were administered the Primary I Battery of the *Metropolitan Achievement Tests.* (The i/t/a edition of the *MAT* was administered to the pupils who were being taught by i/t/a.) In analyzing the test results, it was discovered that the i/t/a group of first grade pupils had obtained significantly higher ($p < .001$) scores in the test areas of Word Knowledge, Word Discrimination, and Reading. In the test area of Arithmetic, the mean raw scores of the i/t/a and the t.o. groups were 52.3 and 52.5 respectively—which the writer viewed as a verification of the initial equality of the two groups.

Three months later, during June, both groups of pupils were administered the regular t.o. edition of the *Stanford Achievement Tests.* Results of this test administration indicated that the i/t/a group of pupils again obtained significantly higher ($p < .01$) scores in the test areas of Word Readiness, Paragraph Meaning, and Vocabulary.

During the last week of school, a noncognitive measure—designed by the writer—was administered to all the first grade pupils. This test was administered orally by the writer with the first grade pupils indicating their answers by circling either "Yes" or "No" on an answer sheet to questions such as:

Do you like numbers better than reading? (No)

Would you like to get a book for a present? (Yes)

For eight of 25 such questions (including the two noted above), the responses of the pupils in the i/t/a group differed significantly ($p < .01$) from the responses of the pupils in the t.o. group.

The second study which is currently being conducted at the elementary level in the Montclair Public Schools is also an experimentally designed study in which an attempt is being made to discover the effectiveness of teaching science by the inquiry method. Fifth and sixth grade pupils will serve as subjects for this study, and these pupils will generally be matched in terms of scholastic aptitude and achievement, and socioeconomic level. The pupils in the experimental group will be taught the science course using Suchman's materials as they are found in the SRA Science Inquiry Development Program. Two of the major problems which may apparently be encountered in this study are: (1) having the classroom teachers teach by the inquiry method; and (2) obtaining an evaluation of the processes rather than the product.

Finally, the writer should mention a research project presently being conducted in the East Orange Public Schools for which he is serving as a consultant. A pilot Intermediate School for fifth and sixth grade pupils has been established in which many innovational features (funded by both ESEA Title III and Ford Foundation) are now in operation. Included among the more recent innovative instructional practices are Computer Assisted Instruction, team teaching and nongraded instruction, and individually prescribed instruction. Other innovative features include open spacing, flexible placement of study carrels, and a comprehensive instructional materials center.

With studies such as those mentioned above being conducted at the present time, research in the teaching and curriculum areas appears to be an increasingly prominent factor in the field of education. And as the findings from this research are implemented, hopefully our school children should be the recipients of better instructional practices in the future.

Symposia

THE DEMONSTRATION OF COMPUTER-AIDED INSTRUCTIONAL SYSTEMS AND OTHER COMPLEX EDUCATIONAL MEDIA

Chairman, Richard L. Wing, BOCES #1, Westchester County, New York

Participants: Robert E. Taylor, Mt. Kisco School District;
Father John Boyle, Fordham University College of Philosophy and Letters; Walter Goodman, BOCES #1, Westchester County.

Father Boyle described the operation and advantages of the closed circuit installation at Loyola Seminary which is designed primarily to train future Jesuit educators in the use of television as an instructional medium. However, Father Boyle stressed, although this is its primary function, the system can serve as a variation of a standard closed circuit origination system that could be incorporated into any school system.

The installation consists of two cameras; one of which is horizontally mounted on a rolling stand that includes lights and monitors, while the other is vertically mounted and fixed into what is called the "Graphics Console."

This installation is unique in that the entire operation can be controlled by one person, the teacher. The teacher controls the switching from one camera to another, and controls, as well, the incorporation of an audio tape recorder and carousel slide projector. The image from the 2 x 2 projector is transmitted under the Graphic Console, and via a multiplex mirror appears on the ground glass screen directly below the vertically mounted camera.

The instructor also determines the composition of the graphic material on the television screen since he can control the zoom lens of the vertically mounted camera. Because all of the controls are within easy reach of the teacher, if he chooses, he can be a teacher, director, technician and cameraman, all in one.

In its very simplicity, this relatively unsophisticated system affords many advantages. First of all, the teacher can be given, in ten minutes, the minimal knowledge he needs to have. A few minor hints about studio techniques completes his orientation. Secondly, the fact that he can come to the studio on his own time, can prepare materials and can practice as long as he likes without the need of gathering a studio crew, opens unlimited opportunities for experimentation in the use of graphics and three dimensional objects.

Experience has shown that the flexibility of the installation has produced a rapid increase in the imaginative creativity of presentations by teachers-in-training.

Father Boyle concluded by saying that closed circuit television could also be used for research.

Mr. Taylor described the Dial Selection System to be in operation at the Middle School of the Bedford Central School System about January 15, 1968.

There are three essential components of any Dial Selection System. First, is the media center, which serves as the master control for the DSS. This area will be equipped with a program distribution and monitoring console, video switching systems, and all program-originating equipment. The console will distribute and monitor programs and will be able when necessary to enable the technician to communicate with persons at any of the stations equipped with dial control plates. Second, carrels, equipped to handle one or two students, will consist of booths with acoustical and physical isolation. Each carrel will provide the student with an opportunity to select any of the available AV materials using the dial control plate or microphone and then the video monitor or headset. Third, teachers in 30 group instruction areas will be able to dial-select or call directly for any of the AV materials distributed from the Media Center. The video and audio portion of the program will be distributed through a monitor mounted on the wall.

Mr. Taylor concluded, by stressing that in spite of the most advanced and sophisticated and complex hardware, it is not the mode of instruction, or the hardware itself that will spell success or failure for Dial Selection Systems; it is rather the software, that is, the programmed materials in the lessons themselves and how they are prepared, selected, and integrated into the curriculum that will be the determining factors. Mr. Goodman stated that since 1962 the Board of Cooperative Educational Services in Bedford Hills, Westchester County, New York, has conducted research on the use of computer-based economic games as an instructional method. The latest project, Project No. 2841, sponsored by Cooperative Research, U.S. Office of Education, is entitled, *The Production and Evaluation of Three Computer-Based Economics Games for the Sixth Grade.*

One of three games developed is the *Sumerian Game,* designed to teach sixth graders some basic principles of economics in operation at the time of the neolithic revolution in Mesopotamia.

The student playing the game, takes the role of a ruler's son in the city-state of Lagash in the year 3500 B.C. He is presented with his first problematic situation: "We have harvested 5,000 bushels of grain to take care of 500 people. How much of this grain shall be set aside for the next season's planting, and how much will be stored in the warehouse? The remainder will be given to the people to eat."

The child makes decisions and enters his answers at the computer terminal. The computer immediately returns a progress report, including the harvest reaped from the seed grain set aside for planting, a word index as to the standard of living, and a report on his inventory. This kind of problem is repeated throughout the entire game, each harvest representing six months in the life of a ruler.

The rule of the first Luduga is devoted to the solution of problems pertaining to an agricultural economy. In the second phase of the game, the child as Luduga II is given the opportunity to apply his surplus grain to the development of crafts. In the third and final stage he is introduced to trade and the more complex problems which confront a changing economy. The rate and trend of development are dependent upon the wisdom of the child's decisions. At intervals the ruler is presented with technological innovations and disasters which will alter his decisions.

The essential idea behind the second game, the *Sierra Leone Development Project,* is to simulate the economic problems of a newly emerging nation, placing a secondary emphasis on a study of the country's culture. The simulated economic and social situations used in this game are drawn from actual problems discussed in the Ten-Year Social and Economic Development Plan now in effect. A specific and already partially successful example is the land reclamation program and the use of this former marshland to cultivate rice in order to diversify agriculture and to decrease expenditures for imported food, in the effort to obtain a favorable balance of trade. As the student, taking a role as an A.I.D. officer, satisfactorily completes and shows an understanding of the problems presented to him, he is promoted to Assistant Affairs Officer and finally to Chief Affairs Officer in A.I.D.

In the third game, called *The Free Enterprise Game,* the player initially assumes the role of the owner of a small toy store. In utilizing the opportunities provided by the simulated store environment and while making decisions to combat those factors in the environment which would be detrimental to the operation of a small business, the player should be able to increase the net worth and monthly sales of his store.

It is hoped that after playing the game the player will realize that each possible decision in a given circumstance tends to carry certain rewards and costs and that the rational decision to make is one which minimizes costs and maximizes rewards within a probabilistic framework.

The educational planning for these games has been accomplished by local teachers and a graduate student at Johns Hopkins.

Equipment used in these projects included IBM 1050 typewriter terminals, an experimental random access slide projector, and an IBM 7090 computer using a time sharing system, so that several pupils could be taught at the same time.

Dr. Wing said that in order to assess the results of the games described previously by Mr. Goodman, an experiment was set up along these lines. From October 1965 to March 1966, 25 sixth grade students from the Mohansic School in Yorktown Heights, New York, played the two games on three terminals at the Center for Educational Services and Research of BOCES. These pupils started in some cases with the *Sumerian Game* and in other cases with the *Sierra Leone Game,* then completing the other. Meanwhile a control class of equal ability studied about the economics of life in Sumer and Sierra Leone under the direction of a talented teacher using ingenious but "conventional" methods.

The basic evaluation problem has been to appraise a combination of two methodological procedures employed in conjunction with each other, one the game method of instruction and the other the use of computer-based typewriter terminals for delivery of the games. The purpose was to find out whether, by using these two methodological approaches, we could provide an individualized learning situation in which the students were able to proceed at their own rates, were able to learn something about economics, and would find the experience stimulating.

The effectiveness of the whole experiment was measured by several different techniques. One was to observe the students carefully and interview them after they finished playing the games. Another was to compare their pretest and posttest scores on specially prepared tests of economic understandings, and a third was a depth interview technique designed to probe for understanding of economic concepts.

The equivalence of the control and experimental groups was established by making a t-test of the difference of mean scores on the California Test of Mental Maturity and the Reading Section of the Iowa Test of Basic Skills, as well as by comparison of scores on the pretest of economic principles. It was found that there was no significant difference between the two groups on these variables.[1]

After a few initial program errors were cleared up, the games ran smoothly. Instruction was self-sufficient to the extent that most of the students required no assistance with the instructional aspects of the games.

In regard to the interest which students had in the games, we found that almost without exception the pupils enjoyed playing these economic games on the computer, and this high interest was maintained throughout the two games, which lasted an average of 15 hours in total.

It was found that in the *Sumerian Game* the experimental group gained a significantly greater amount from pretest to posttest than the control group. This difference in gain was significant at the 1 percent level of confidence. In the *Sierra Leone Game* the control group scored slightly but not significantly better than the experimental group.

It appears from our testing that we should not claim that the computer-game technique is clearly superior to classroom instruction. It does seem, however, that the experimental method is at least as effective, to judge by the measurement which we employed.

In the experimental setting the average pupil time on the *Sumerian Game* was 10 hours and 15 minutes with a range from 6 hours and 40 minutes to 14 hours and 5 minutes. The average time on the *Sierra Leone Game* was 5 hours and 5 minutes with a range from 3 hours and 55 minutes to 8 hours and 0 minutes. The combined average then is about fifteen and a half hours to complete both games.

It seems, therefore, that studying by games on the computer took less time than studying the same material in class.

Relationships Between Learning and Factors of Individual Difference

Table of Correlations

Gain, Pretest—Posttest versus	Sumerian Game	Sierra Leone Game
a. Time on Computer	−.199	−.374
b. Reading Ability	+.228	+.417
c. Intelligence	+.388	+.546

a) Apparently the students who spent the least time at the computer made the greatest gains.

b) There was a low positive correlation between reading ability and improvement of test score.

c) The more intelligent students gained more from pretest to posttest than the less intelligent students, even though there is a ceiling on the test which prevents the best students from displaying the whole range of their talents.

General conclusion: Analysis of the computer-based economics games project demonstrated that there is sufficient promise in this kind of instruction to warrant further, exploratory research both with the computer medium and the game mode.

---

[1] Francis G. Cornell is project statistician.

NEEDED RESEARCH IN TEACHER EDUCATION: THEORY BASED PROGRAMS AND TEACHER EDUCATION

Lawrence V. Castiglione, City University of New York, Queens College

Systematic, comprehensive theories of teaching, in the formal sense of the term, do not now exist. What should be the nature and function of teacher education programs is unclear (Anderson, 1962). Glazer (1966) feels that statements of educational goals are platitudes, subject to so many interpretations that evaluation in terms of specific student behaviors is impossible. Only recently have researchers begun to focus their attention on the teaching process. It is hardly surprising, given this state of affairs, that there are not programs of teacher education that have implicit in them a theory or theories of teaching. The two necessary, if not sufficient conditions—a clear statement of fundamental policy and an empirically grounded, comprehensive theory of teaching—have not been met.

Scandura (1966) along with others, maintains that at present teaching is a practical endeavor, and its practice more an art than a science. He holds that there are certain utilities that are characteristic of science and technology that are not shared by an art, and that education would benefit by the transition from art to scientific technology. Skinner claimed that education is in principle a scientific technology that could be made operational if cultural inertia were overcome. Skinner expressed this opinion in 1954. To date there is no substantial evidence that teacher education has become or is about to become a scientific technology.

The delay in implementing what is already known may be in part due to the fact that the concerns of classroom teachers overlap only slightly the concerns of psychologists in laboratories. This is most apparent in the psychological study of learning. Bruner (1964) writes that learning theories focus upon a description of the processes of learning, and that teaching theories, as distinguished from learning theories, should be concerned with specifying the means by which learning may be facilitated. Ausabel (1966) takes the position that this distinction holds in a formal sense only, in that theories of learning, by specifying optimal conditions and the variables affecting learning, imply appropriate teacher behavior.

However, there is little evidence to support the notion that complex molecular systems, however clearly described to a student, will result in appropriate behavior in the absence of detailed instruction directed to that end. Travers (1965, p. 529) states that, "Teachers do not change their ways of behaving simply by being told that learning would proceed with greater efficiency if they behaved differently." James Thurber (Mandler & Kesson, 1964) once wrote that "A word to the wise is not sufficient if it doesn't make any sense."

Perhaps the difficulty in communication between researchers and teachers might be reduced by employing instructional practices suggested by learning theories to the training of teachers, rather than implying, as Schutz and Silberman (1966, p. 298) put it, ". . . that the process of extrapolating the theory into practice is simple and straightforward." For example, developing and using techniques for simulating specific aspects of classroom teaching operations in order to develop proficiency, applying instructional devices and materials, and using diagnostic evaluations might be as effective as lecturing.

Presumably teaching theories would permit one to prescribe behaviors that would facilitate learning in a predictable manner under known conditions. Teachers should be able to decide between various alternatives on a systematic basis within the context of a standard classroom setting.

Taxonomies of teacher behavior derived from already existing observational systems might conceivably form the basis, or elements, of a theory of teaching. The classes of these observational systems range from power relationships (Hughes, 1963) to classifications borrowed from information theory (Ryans, 1963). Scandura (1966) points out that beyond Gage's suggestion that teaching be treated as an independent variable and learning as a dependent variable, teaching theories must take into account individual differences in achievement and ability as well as the timing and sequencing of information.

Groen and Atkinson (1966) have proposed a model for the dynamic programming of a class of instructional problems that treats instruction as a multistage decision process which is subject to evaluation by means of a mathematical theory. Optimal stimulus presentation strategies, sensitive or insensitive to students' responses, may be approximated. Although not directly applicable to classroom teaching, this model may provide a useful conceptual approach in formulating a theory of teaching. Certainly, the idea of stimulus presentation decisions as an instructional strategy directed toward clearly defined objectives of instruction, may serve as one possible locus of transition from the art of teaching to the science of instruction.

It is obviously necessary to have a well developed theory of teaching to enable one to clearly specify important elements of behavior, to formulate alternative hypotheses, and to meaningfully interpret differences between experimental treatments. Research should attempt to identify the major sources of variance in student learning attributable to teacher behaviors, and the influence on school learning of these behaviors under classroom conditions. The current lack of comprehensive theory notwithstanding, even a pre-theoretical model with limited application would establish a basis for evaluating the effectiveness of instruction interpretable within the context of the model employed. An attempt in this direction, almost by definition, would move educational practice and programs closer to becoming a scientific technology than has thus far been the case.

Bibliography

Anderson, A. W. The teaching profession: an example of diversity in training and function. In N. B. Henry (Ed.), *Education for the professions; 61st Yearb. Nat. Soc. Stud. Educ., part II.* Chicago: U. of Chicago Press, 1962. Pp. 140-67.

Ausabel, D. P. Book review, *Harvard educ. Rev.*, 1966, *36*, 337-40.

Bruner, J. S. *Toward a theory of instruction.* Cambridge, Mass.: University Press, 1964.

Glazer, R. Discussion: the effect of educational research on classroom instruction. *Harvard Educ. Rev.*, *36*, 315-17.

Groen, G. J. & Atkinson, R. C. Models for optimizing the learning process. *Psych. Bull.* 1966, *4*, 309-20.

Hughes, Marie M. Utah study of the assessment of teaching. In A. Bellack (Ed.), *Theory and research in teaching.* New York: Teachers Coll., Columbia U., 1963. Pp. 25-36.

Ryans, D. G. Teacher behavior theory and research: implications for teacher education. *J. of Teacher Educ.*, 1963, *14*, 274-93.

Scandura, J. M. Teaching—technology or theory. *Amer. educ. res. J.*, 1966, *2*, 139-46.

Schutz, R. E. & Silberman, H. F. Discussion: the effect of educational research on classroom instruction. *Harvard educ. Rev.*, 1966, *36*, 297-301.

Skinner, B. F. The science of learning and the art of teaching. *Harvard educ. Rev.*, 1954, *24*, 86-97.

Thurber, J. Quoted by G. Mandler & W. Kesson in *The language of psychology,* New York: Wiley, 1964.

Travers, R. M. W. A study of the relationship of psychological research to educational practice. In R. Glazer, ed., *Training research and education.* New York: Wiley Science Editions, 1965. Pp. 525-58.

NEEDED RESEARCH IN TEACHER EDUCATION: SENSITIVITY TRAINING AND THE PROCESS OF CHANGE

Julian Roberts, Yeshiva University

One important aspect of teacher education that needs further study is the preparation of preservice teachers for adequate handling of the phenomenon of change. The several aspects of teacher education that emphasize change in the curriculum, change in role perception of the teacher and of the learner, change in the notions of how instruction is truly individualized, somehow have not really explored how newcomers to the profession may enter a world where these aforementioned changes may serve to threaten individuals, or even whole school systems.

In the time allotted, I would like to share with you one illustration of the processes involved in introducing change into a school setting, simply to highlight the complexity of the problem. In a suburban community on Long Island, a pilot project was initiated last year to explore the possibility of changing the noted lack of effective communications among the several levels of the community. Four adult community members, representing diverse roles (parent, ex-school board president, the leaders of the 2 forces pro and con in a proposed bond issue), the school superintendent, 3 teachers, and 7 eleventh grade students engaged in a t-group experience. T-group is a process developed by the National Training Laboratories, a subsidiary of the NEA. It is defined by Benne and others as follows:

> A T-group is a relatively unstructured group in which individuals participate as learners. The data for learning are not outside these individuals or remote from their immediate experience within the T-group. The data are the transactions among members, their own behavior in the group, as they struggle to create a productive and viable organization, a miniature society; and as they work to stimulate and support one another's learning within that society. Involving experiences are a necessary, but not the only, condition of learning. T-group members must establish a process of inquiry in which data about their own behaviors are collected and analyzed simultaneously with the experience which generates the behavior. Learnings thus achieved are tested and generalized for continuing use. Each individual may learn about his own motives, feelings and strategies in dealing with other persons. He learns also of the reactions he produces in others as he interacts with them. From the confrontation of intentions and effects, he locates barriers to full and autonomous functioning in his relations with others. Out of these he develops new images of potentiality in himself and seeks help from others in converting potentialities into actualities.
>
> Each individual may learn also about groups in the processes of helping to build one. He may develop skills of membership and skills for changing and improving his social environment as well as himself. The staff who work with T-groups do not see any necessary opposition between participation in groups and autonomous individual functioning, though they are well aware that opposition does occur in many associations of our lives and that group forces may be used to inhibit personal development. In the T-group, on the contrary, the objective is to mobilize group forces to support the growth of members as unique individuals simultaneously with their growth as collaborators. Influences among peers are paramount in this learning process. In the T-group, members develop their own skills in giving and receiving help. They learn to help the trainer or teacher as he assists in the development of individual and group learnings.[1]

The National Training Lab lists the following five factors as most important objectives of human relations sensitivity training:

Self-insight

Better understanding of other persons and awareness of one's impact on them.

Better understanding of group processes and increased skill in achieving group effectiveness.

Increased recognition of the characteristics of larger social systems.

Greater awareness of the dynamics of change.

A training laboratory tries to create a climate encouraging learnings, understandings, insights, skills in the areas of self, group, and organization. The preliminary examination of the responses of the participants to two questionnaires indicates that a notable degree of achievement of these objectives has been reached.

---

[1] Leland Bradford, Jack R. Gibb, and Kenneth D. Benne, *T-Group Theory and Laboratory Method.* New York; John Wiley & Sons, Inc., 1964.

An analysis of the process of T-group reveals its relationship to the total complex of the process of change. The rationale indicates that before change can take place effectively, people affected by such change must learn how to interact with each other so that greater productivity may result as the change takes effect. From this experience, a present research project was evolved and is now in operation. How can a classroom setting be changed so that interaction between the student and teacher, and student and student may give everyone an opportunity to provide maximum performance of all involved in the educational process?

Such a change necessitates new role concepts for both teacher and student. The teacher must move from a highly directive person and leader to one who becomes a more highly integrative leader, one who readily accepts statements of choice, differences of opinion, and suggestions for new ways of proceeding within the classroom setting. The student too must move from one who seeks to develop thinking patterns convergent with those of the teacher or those he feels are the pacesetters of his peer group to one who is not afraid to express divergent thinking. His role becomes more creative, less limited in response pattern, more original. He, too, must exhibit behavior that recognizes differences, and willingness to accept responsibility for suggesting new ways of proceeding.

Changes such as those suggested above may threaten both teachers and students. T-group training in self-awareness, in ability to listen to others, to see how others perceive you and how you rightly or wrongly have perceived others is one possible way to handle the threats posed by change. The classroom setting used in this experiment is 12th grade social studies: Problems of Democracy. Hilda Taba points out that social studies are always involved with feelings and attitudes. Thus a class in Problems of Democracy must be concerned with notions of self, of uniqueness, of group membership, and group process. The climate to be developed in such a class must include: a) the capacity to identify with people from different backgrounds within the community, avoiding or overcoming stereotypes that hinder real understanding of, and respect for others; b) the self-security that permits one to be comfortable in differing from others; c) the openmindedness that examines the opinions and ways of others with reasonable consideration and objectivity; d) the acceptance of changes that adjusts as a matter of course to new ways and events; e) the tolerance of uncertainty and ambiguity without anxiety; and f) the responsivness to democratic and human values.[2]

T-group should allow for a greater degree of achievement of these objectives. The very essence of sensitivity training calls for the development of personal attributes and abilities to cope with elements in the environment that are inhibiting the teaching-learning process.

What this short paper is suggesting, then, is that further research in the nature and process of change is mandatory if climates that will promote change in the most positive and profitable sense are to be established in the school setting. The work of Bellak on the language of the classroom has already indicated the degree to which the oral behavior of teacher to student or student to student may inhibit an individual's attempt to recognize, use and develop his potential, and his ability to learn. Other techniques must be investigated. If we mean to produce teachers who will be prepared to do more than perpetuate an educational system that has revealed its shortcomings, then Federal, State, and local support must be given to encourage further research in this area.

---

[2] Hilda Taba and James L. Hills. *Teacher Handbook for Contra Costa Social Studies Grades 1-6,* San Francisco, California 1965.

CHARACTERISTICS OF NEW ENGLAND CATHOLIC SCHOOL PRINCIPALS

Chairman, John A. Schmitt, Boston College

That Catholic parochial schools have problems distinguishing them from their public school counterparts is nothing new to persons familiar with parochial school systems. The fact may be surprising, however, to those who have no more than a surface familiarity with the system. The symposium participants whose comments comprised this presentation are intimately involved with the system; they understand the problems; they have determined the facts of the situation, and they have elicited significant conclusions.

Available space does not permit presentation of the findings in detail.[1] However, the procedural description by Father O'Neill details the procedures employed, and the commentaries by Doctor Nuccio and Doctor Madaus illustrate the major implications of the findings.

Father McHugh's introduction includes a general description of the objectives of the New England Catholic Education Center (NECEC). Notwithstanding ample research evidence concerning the elements that distinguish parochial schools from public schools, relatively little effort has been devoted to determining the needs which are peculiar to parochial schools. NECEC strives to reduce this particular knowledge vacuum by supporting research oriented, specifically, toward Catholic parochial schools, and this symposium reports one example of such an effort.

NECEC AND RESEARCH IN CATHOLIC EDUCATION

Reverend Paul F. McHugh, New England Catholic Education Center

These remarks attempt to establish an understandable framework within which this research took place and which, building upon the empirical evidence reported, will provide the support for current and anticipated research. This research framework is an embryonic development and service center known as NECEC, the New England Catholic Education Center.

There are nearly one-half million children of elementary and secondary school age attending 900 elementary and 300 secondary Catholic schools, staffed by 16,000 religious and lay teachers. Daily administrative and curriculum decisions concerning these children are being made in administrative patterns which cluster around eleven diocesan superintendents of schools, who are the chief school administrative officers of these eleven ecclesiastically jurisdictional areas.

The hallmark of these decisions has been the historically conspicuous absence of data to verify and support such important decisions. NECEC is the institution, created at Boston College and utilizing the total resources of that university, which is attempting to conduct research consonant with the needs of New England and pointed toward constructive and continuing change.

Hypothesizing the school principal as the educational leader, pivotal in effecting change, the Center's first research activities concerned themselves with research about the principals in approximately 1,200 New England Catholic elementary and secondary schools. Immediate funds were allocated to three studies:

1) The role and function of elementary school principals, conducted by Rev. Patrick O'Neill
2) A similar study of the secondary school principalship, conducted by Mr. Thomas Whalen
3) As a further secondary school refinement, a study of the staff perceptions of the role and functions of the principal as educational leader in randomly selected secondary schools.

The findings and implications of the first two of these studies are reported in this symposium.

It is well to add, that aside from presenting the results of these studies, it is our desire to convey the idea that NECEC represents a concrete proposal to alleviate a longstanding gap between needed research to satisfy the requirements of school administrators and the research (and the researchers') need for administrative support—concrete incubator settings for research, field testing, and dissemination of research findings.

---

[1] The documents reporting the general findings reported by this symposium consumed more than a thousand pages in four volumes. Inquiries regarding specific findings may be directed to: New England Catholic Education Center, Boston College, Chestnut Hill, Massachusetts 02167.

THE STUDIES OF CATHOLIC SCHOOL PRINCIPALS' CHARACTERISTICS

Reverend Patrick J. O'Neill, Diocese of Fall River, Massachusetts

The studies conducted by Thomas Whalen and me were concerned, respectively, with the principals of Catholic secondary and elementary schools in the eleven dioceses of New England. Data were gathered by means of a questionnaire which sought information regarding the principals' personal characteristics, the characteristics of their schools, the activities in which the principals engaged, their evaluation of the importance of these activities, their opinion regarding the quality of their own preparation for the various activities, and opinions regarding a sample of educational problems and particular needs of principals.

The questionnaire, accompanied by a letter from NECEC's Executive Director, was mailed to every principal of a Catholic school in New England. The letter explained the purposes of the studies, assured the principals of anonymity, and solicited their cooperation. Followup letters were directed to those principals whose questionnaires had not been returned within one month, and later telephone followups were made to those still failing to respond. As a result of these measures, responses were obtained from 771 elementary school principals (87%) and 205 secondary school principals (83%).

Questionnaires were precoded with an identification number to identify each school, and each questionnaire item was identified by a number indicating the card column into which the item response was to be punched. Transference of the data to punched cards was thereby greatly facilitated, and the data were next recorded on magnetic tape for computer processing.

Responses to objective questions on the questionnaires were tabulated separately for each of the eleven dioceses, as well as for the total group. Percentages were calculated, and chi square tests were employed to identify items which distinguished individual dioceses from one another. Relationships between certain of the principals' personal characteristics and their various activities, opinions, and attitudes were also tested for statistical significance.

While space does not permit complete reporting of the findings, the following general conclusions were considered by the investigators to be of particular significance:

1) The role of the principal, in relation to the parish, the diocese and his religious community—the vast majority of principals being members of religious communities—is very poorly defined
2) The principalship, for Catholic schools, has not been developed as a distinct profession with a definite process of selection, special preparation and sufficient term of office
3) The principals, in many instances, have not been taking advantage of the leadership opportunities presently available to them

The studies which have been briefly described here have most serious implications. Clarification of principals' roles in relation to pastors, diocesan superintendents and their own communities would appear to be essential, if the effectiveness of the principal is to be increased. Potential principals should be identified early in their careers as members of religious communities through objective processes; they should receive appropriate administrative training; and their tenure as principals should be assured to maximize the return upon the investment in their training. For those presently occupying administrative positions, there is a great need for inservice training in administrative principles, supervision, human relations, and curriculum development. Freedom from teaching responsibilities, a modest amount of secretarial assistance, and active participation in the budget-making process relating to their schools are additional, imperative needs.

THE NECEC PRINCIPALSHIP STUDIES AND ADMINISTRATIVE THEORY

George F. Madaus, New England Catholic Education Center

The O'Neill and Whalen studies pose some interesting questions in the light of administrative theory. Space permits consideration of only two aspects of the relationships between these data and the theory of Getzels and Coladarci.

In theory, a Catholic school principal can base his authority on one of three legitimate sources: traditional, charismatic, or rational. Traditional authority is based on established belief in the sanctity of the traditions under which the authority figure exercises authority, and the leader is obeyed because he occupies a traditionally sanctioned position. Charisma derives from the exemplary character or general wisdom of the leader, whom the group obeys because of their personal trust in him and the normative patterns ordained by him. Pacifist attraction to Bertrand Russell, because of his general wisdom, and the personal trust in the leader and his normative patterns demonstrated by early members of the Mormon Church illustrate charismatic authority.

Both studies found that the role of school principal was very often combined with that of local community superior, wherein the superior draws on both traditional and charismatic sources for authority which may transfer to the principal role. However, the principal should operate from effective authority which originates with the group and is given to the leader by the group. Group members can withhold it without depriving the administrator of his formal position, because he retains the power to command and direct.

Teachers and principal are members of the educational profession, and each is expert in his specialty within that field. The principal specializes in the allocation and integration of human and material resources among tasks, so that the goals of the institution may be achieved.

Getzels and Coladarci's theory holds that school administrators trying to attain authority without technical competence in their specialty are doomed to failure, because they weaken their authority each time they appeal to traditional or charismatic sources. Their authority is reinforced and strengthened, however, each time decisions are based on technical competence.

The present studies show that the Catholic school administrator has little past experience and formal training in educational administration. There may be a tendency, therefore, to try to base authority on traditional or charismatic rather than effective grounds. Determination of the sources of authority upon which the Catholic school principal relies will require further study.

The area of role conflict, alluded to in the discussion of sources of authority, is also of special interest. Educational roles can be functionally specific or diffuse, and specificity describes the *range* and *scope* of roles and facilities that may *legitimately* be included as matters for allocation and integration within a given interaction. If the range of roles is specific, the patterns of interaction are strictly limited in scope of interest. Interaction and obligations are restricted to those roles defined by the technical competence and status of the participants.

The findings that principals felt they could not adequately handle the dual role of community superior and school principal, documents the community diffuseness in range of roles that infiltrate religious teaching. Religious responsibilities, tangential to the teaching role, can become confused, and the principal may find it difficult to make decisions on the basis of technical competence, thus weakening effective authority and forcing a reliance on traditional authority.

An investigation of the effects of role diffuseness on administrative effectiveness would certainly be worthwhile, shedding light on assumptions of the theory and affording data for decisions regarding a combination of the two roles. Further exploration of these studies in the light of administrative theory could improve the theory, while contributing substantially to the art of administration in Catholic schools.

IMPLICATIONS OF THE STUDIES FOR HIGHER EDUCATION

Vincent C. Nuccio, Boston College

These studies and others in which Boston College is presently involved with NECEC emphasize three basic needs: first, to establish on university campuses a climate for cooperative programs involving nonpublic school personnel, particularly administrators. Secondly, to involve nonpublic school pupils, teachers and administrators (to which Catholic schools contribute large numbers) in future educational research projects. Third, the need for universities to present in-service programs and preparation programs which can upgrade and prepare nonpublic school personnel.

The fulfillment of these needs will require a better understanding of, and a closer relationship with, those who are responsible for the identification and preparation of nonpublic school personnel. In the case of Catholic schools, this would mean a working relationship with the religious superiors of communities presently involved in education.

Nonpublic education is changing as a result of recent Federal legislation, and this has very definite implications for the future preparation of nonpublic school elementary and secondary principals. The University Council on Educational Administration (UCEA) has recognized this need and is proposing, to the Kellogg Foundation, a consortium of the University of Chicago, Loyola University of Chicago, Michigan State University, St. Louis University, and the University of Texas. It also has proposed a consortium within the consortium of Harvard University, Boston College, NECEC, and the National Association of Independent Schools. This UCEA project recognizes ". . . the need for more effective research, development, training and dissemination programs pertinent to non-public school administration in both its preparatory and practical aspects."

More institutions of higher education need to get involved in programs similar to the model of NECEC at Boston College, where a regional organization is established for the purpose of research and dissemination aimed specifically at Catholic schools. A number of these university-affiliated centers might comprise a national network of centers for the study of Catholic education, and such a network could considerably enhance the prospects for future research in Catholic education. Planning, educational programming, and a system of information retrieval would substantially improve the prospects of assembling sound information on which to base decisions concerning the future of Catholic schools.

TAXONOMIC RESEARCH IN ORGANIZATIONAL BEHAVIOR IN EDUCATIONAL SYSTEMS

Frank W. Lutz, Mozell Hill, Paul Cullinan, & Seymour Evans, New York University;

Glenn Immegart, University of Rochester, & Harold Robbins, Queens College

Introduction

The project "Developing Taxonomies of Organizational Behavior in Education" was undertaken by New York University under direction of Dean Daniel E. Griffiths.[1] It was funded by the U.S. Office of Education with the cooperation of the University Council for Educational Administration and utilized researchers from five different institutions. The purposes of the project were several: 1) To define conceptual boundaries of organizational behavior in education; 2) To provide guides for abstracting content on educational administration; 3) To furnish guides in developing instructional materials; 4) To encourage and provide guides for the synthesizing of knowledge; and 5) To stimulate future research.

The research was undertaken in the following steps and will be presented here in that manner. An eight month field study in an educational organization was undertaken in order to gather data. At the same time, four theories were selected from which original classification characteristics were developed, and then modified and tested with the field study data. These theories were: general systems, bureaucracy, compliance (not covered in this presentation[2]), and decision making. The final step in the project was an attempt to synthesize the taxonomies into a single taxonomy.

Purposes and Processes of Taxonomic Inquiry

The purpose of taxonomic inquiry is to obtain an arrangement or ordering of phenomena into a readily understandable classification system according to evidence supporting theoretical relationships. To achieve the best possible approximation to this organization, the process of inquiry depends upon theories, principles and procedures that are, strictly speaking, the "science" of taxonomy. Formulating and testing theories and applying these principles are the most basic things a taxonomist does. But he does them by means of a set of procedures such as those in Table 1.

In taxonomic inquiry, the most critical steps in the series are those of synthesizing relationships, arranging similarity matrices, and applying techniques of cluster analysis. These three operations are concerned with the conceptual aggregation of phenomena into classes and the arrangement of classes into a hierarchical system.

1. *Purposes and criteria.* The major purpose of taxonomic inquiry is to construct classes about which generalizations can be made. These classes are developed in accordance with a purpose by means of theories which yield classification systems for a certain range of phenomena. Although a number of classification schemes may be possible, one *form* of classification should be applied to all phenomena with which the taxonomy is concerned.

2. *Conceptual bases.* In taxonomic inquiry classifications are conceived as theoretical entities, which operationally specified, are explicable in terms of categorical logic and utilize statistical methods.

3. *Principles and procedures.* Classes or aggregations of phenomena are the basic units of taxonomy. Taxonomic studies are always statistical in nature which provides evidence that the theoretical specifications of classes are met. The classes at all levels of a taxonomy are defined by their relationships expressed in conceptual terms. The most significant relationships, consistent with theory, guide the development of classifications and hierarchical ordering.

The principles and procedures outlined here have served to guide the empirical processes in the investigation of organizational behavior in education. However, numerous problems occur no matter how tractable the phenomena are for taxonomic inquiry.

The Field Study

Based upon the belief that organizational behavior centers around organizational roles, six roles were elected as focal points of observation within the school district: The School Board, The Superintendent, The Elementary Curriculum Director, The Senior High

---

[1] This project was funded by the U.S. Office of Education, Contract #H5-0792-2-12-1.

[2] Dean Stephen Hencley developed this schema but was unable to be present in Albany for the ERANYS meeting.

TABLE I THE FUNCTIONS OF TAXONOMIC INQUIRY INTO ORGANIZATIONAL BEHAVIOR IN EDUCATION

Functions	Procedures	Purposes
Observing and Recording	Selecting representative field of study Devising methods of observation Determining criteria for unit of behavior Preparing transcripts of observations	Depict organizational behavior in education Discriminate patterns of behavior Identify operational units Record tractable specimens
Formulating Conceptual Schemata	Stating propositions in terms of organizational behavior in education Indicating categories of observational predictiveness Stating parameters of categories of schemata Assigning characteristics to observational protocols	Explicate conceptual frameworks Operationally specify classes Characterize classes of observables Apply scheme to recorded specimens
Classifying Behavior Units	Preparing frequency distribution charts Checking frequencies versus proposed categories Analyzing by various statistical methods Indicating coefficients of association	Compare and contrast analyzed specimens Corroborate theoretical referents Examine degrees of affinity Synthesize relationships
Developing Taxonomies	Arranging similarity matrices Applying techniques of cluster analysis Using criteria of rank ordering Modifying categories; initiating new study	Construct classes from data Evaluate polythetic features of divisions Assemble various taxonomic groups Refine and test revised schemata

School Principal, The Junior High School Principal, and An Elementary School Principal.

Ground rules had to be adopted in order to delimit the parameters of what was termed a single unit or specimen of organizational behavior. The unit was described as the occurrence of a topic of behavior in the life of a central actor. The unit began when the topic was introduced and it continued to be recorded until that topic was terminated and another topic was introduced. Even though the same topic may have been initiated again at a later time, the termination at that given point ended the unit. It was decided to define this O.T.U. (organizational taxonomic unit) as a unit comprised of the following qualities: (1) the purpose of the behavior had to be identifiable; (2) there had to be identifiable actors; (3) there had to be some history leading to the unit of activity; (4) there had to be a description of what happened during the elapsed time interval of observation; and (5) it had to be possible to state what happened as a result of the activity. It was also decided to include actual dialogue which occurred in order to supply detail to the unit. Other suggested methods of defining a unit of behavior did not permit all taxonomists to classify all units. The O.T.U. defined the occurrence of behavior which the taxonomists were then required to classify. Ninety O.T.U.'s were developed from the collected descriptions of behavior reported in the extensive field notes. Each taxonomist was able to focus upon the same descriptions of behavior.

Seven, separate, previously trained observers worked in the field to collect the data concerning the sentiments exchanged during interactional activity of the individuals within the school district. Having agreed that the Homanian concepts of activity, interaction, and sentiment were basic elements of human behavior, the task of reporting the happenings was refined to insure that a reliable description of the behavior was made available to the taxonomists. Also, this method of data collection mode did not then dictate the framework of the subsequent analysis.

It is not possible to observe and record the totality of behavior in any given milieu. However, the researcher can describe his instrument, and the object of his observations, record what he observes in behavioral terms, and tell how he has checked the reliability of his observations. Although no two observers were assigned to the same activity during the course of the field collection, observations were seen to overlap in many ways. Related material, transcribed from the notes of different observers, provided another means of checking upon the reliability of the data collection.

Sound, motion picture films and magnetic tapes made in conjunction with the narrative reporting, afforded opportunity for third party review. In addition, the amount of recorded, observed units numbered in the hundreds, thus establishing some increase in the internal consistency of the method.

The selection of the O.T.U.'s from the total behavioral sequence in an educational organization could be arbitrary. In order, better to avoid this, a theoretical grid was developed to make more certain that samples could be obtained which represented all the organizational levels in the school as well as all the role incumbencies thought to be important to the classification schema.

Neither the field study nor the O.T.U.'s are intended to be descriptive of the universe of organizational behavior. The goal of the study was not the development of an image of a "typical" school enterprise, rather, it was the purpose of the field study to present a picture of a single school district.

General Systems Based Schema

Systems theory was selected as an approach to taxonomic research into organizational behavior in education as a result of its demonstrated value as a theoretical construct, as a vehicle for understanding organizational phenomena, and as a classificatory device. Since no one universally accepted or all inclusive "systems theory" exists, a four scheme approach to classification was formulated using four distinct theoretical perspectives for conceptualizing systems.

The schemes developed along with their theoretical referents, classification procedures, and functions are indicatively described below:

1. CLASSIFICATION APPROACH I: *Comprehensive Systems Characterization Scheme.*
   Derivation: from "systems theories of the whole" or comprehensive systems theories.
   Classification Categories: actors, inputs, organizational mechanism (s), organizational subsystem (s), outputs, and locus of loci of forces.
   Method of Characterizing Specimens: verbal description relative to the above categories of each unit of organizational behavior (specimen).
   Function: (1) to organize each behavioral unit or specimen into an orderly, workable "whole" in terms of basic systems terminology and concepts; and (2) to reduce data relative to each behavioral specimen into a form amenable to inspection or content analysis for classification purposes.
2. CLASSIFICATION APPROACH II: *Input-Output Linkage (Subsystems) Scheme.*
   Derivation: from process or subsystem theories using standard organizational terminology and standard educational task areas and terminology to provide specific operational subcategories within the framework of the process or subsystem theories.
   Classification Categories: inputs in the form of operands (information, energy, and resources) and operators (control structures, operations, and personnel); functional subsystems of the following types—administrative, supervisory, instructional, purpose determination, personnel, recordkeeping, client-constituent relations, business management, and negotiation; and output in the form of productivity (products or performance), affectivity, and feedback (internal and external).
   Method of Characterizing Specimens: checking the presence of detailed attributes (subdivisions of the above generic classification categories) on a structured worksheet.
   Function: to microscopically characterize behavioral specimens in terms of (1) inputs, subsystems, and outputs; and (2) input-output linkage through functional subsystems.
3. CLASSIFICATION APPROACH III: *Analysis of System Properties—States and Processes.*
   Derivation: from theories of universal open system properties.
   Classification Categories: input, output, input-output relationship, steady state, relationship of functional subsystems, self-regulation, feedback, negentropy, progressive segregation, progressive mechanization, and equilibrium.
   Method of Characterizing Specimens: checking appropriate detailed characteristics under each category on a structured worksheet.
   Function: to macroscopically characterize behavioral specimens at a given point in time-space in terms of the universal properties exhibited by all open systems.
4. CLASSIFICATION APPROACH IV: *Output Analysis Scheme.*
   Derivation: from output theories or system outcome analysis.
   Classification Categories: productivity (in terms of product and services utility), organizational health (in terms of adaptability, identity sense, and reality test capacity), integration potential (in terms of self-actualization, group decision-making, and individual change flexibility), and feedback (in terms of desirability and penetration).
   Method of Characterizing Specimens: rating of each of the classification categories on a four-point scale relative to degree of presence (from low to high) and plotting profiles for each specimen.
   Function: (1) to qualitatively assess system output variables and system achievement; and (2) to indicate the degree of system "openness."

After ninety units of organizational behavior in education were classified according to each of the schemes, data analysis was undertaken to ascertain (1) whether the devised schemes sorted the objects classified into taxonomic groupings, and (2) whether any such groupings revealed facilitate the understandings of the objects or phenomena so ordered.

As a result of preliminary classification category use and overlap analysis it was found that: (1) the domain covered by the behavioral units classified could be described in "systems" terms; (2) the classification schemes contained no discernible redundancy; (3) the devised schemes evinced a satisfactory level of discriminatory power; (4) one of the schemes (the first) was less useful for ultimate purposes of the study than the others; (5) more rigorous analyses using established taxonomic approaches could be profitably pursued.

Additional analyses of scheme II, III, and IV classification data were undertaken using both the numerical and theoretical taxonomic approaches. As a result it was concluded that:

1. The devised schemes do provide a basis for sorting the classified units of organizational behavior in education into taxonomic groupings (e.g. species).
2. The schemes permit the logical derivation of sub (e.g. sub-species) and supra (e.g. archetypes) taxonomic groupings.
3. The "essence" of each of these taxonomic groupings is describable in terms of systems concepts which differ uniquely from existing organizational terminology and nomenclatures.
4. The taxonomic groupings and relationships revealed have a potential for facilitating the understanding of organizational behavior in education. (In this regard a number of leads for further research including the generation of specific research hypotheses, resulted from the study.)
5. Application of the classification schemes to a larger behavioral unit sample, drawing on diverse kinds of educational organizations, along with more extensive data analysis appears to be a promising direction for taxonomic research into organizational behavior in education.

Decision-Making Based Classification Schema

The classification system developed by Griffiths utilizes one hundred forty-two characteristics arranged in an ordered relationship which account for the range of behaviors that must be considered in the description of administrators at work. Rather than attempt to describe each category and the relationships among them, the three broad components under which they are subsumed will be outlined. This approach enables one to quickly grasp the conceptual framework used; this schema can be illustrated as follows:

Input
↓
Decision-Process
↓
Output

Each O.T.U. had an input; that is, some information stimuli received by the central actor generated a problem or an activity which resulted in interaction between or among actors. If a problem was present in the O.T.U., some type of decision-making followed. This behavior resulted in a judgment which affected a course of action and ranged from the postponement of a decision to the making of a final decision.

Other behaviors reported appeared to be pointless; that is, the actors engaged non-decision activities. Such behaviors were not organizationally task oriented and not concerned with maintaining relations. While at times information was exchanged, such behavior was not purposive and could best be described as the mere expenditure of energy. Regardless of whether a decision-making process occurred, there was always, or so it seemed, an outcome present. It may have been precisely this reason; that is the frequent occurrence in the O.T.U.s of administrative behavior obviously unrelated to decision-making that led Griffiths to the final classification system.

Based on certain assumptions regarding the purposes of administrative behavior, the theory, using a set of decision-making concepts, attempts to explain administrative behavior. Three assumptions are:

1. *Administration is a generalized type of behavior which is found in all human organizations.*
2. *Administration is the process of directing and controlling life in a social organization.*
3. *The specific function of administration is to develop and regulate the decision-making process in the most effective manner possible.*

This theory views the process of administration as the monitoring of decisions made in the organization and holds that administrative behavior can be described and understood in terms of decision-making or the monitoring of decisions in the organization.

Griffiths then formulated the steps which were descriptive of the decision-making process. These steps, which in effect was an "armchair" version of the theory were further refined in the study of the performance of elementary principals in a simulated situation.[3] It was found that the scoring categories loaded heavily on one of the two secondary factors. Factor X, *Preparation for Decision vs. Taking Final Action,* was very similar to the theoretically conceived steps used to describe the decision-making process.

Drawing heavily from the simulation study and other concepts derived from the expanded theory, classification categories were developed. In scoring the O.T.U.s it became immediately apparent that the categories were inadequate. Much of the administrative behavior reported in the O.T.U.s was obviously unrelated to decision-making and it was decided that this classification schema should be abandoned.

Returning to the situation study, it was found that the other of the secondary factors, Factor Y, *Amount of Work Expended in Handling the Item,* could be applied to account for much of the nondecision-making behavior reported in the O.T.U.s. While Factor X focuses on decision-making behavior, Factor Y captures much that goes on in an organization which can best be described as the expenditure of energy, work, and the like. The cues provided by Factor Y led to an expanded classification schema which, although neither mathematically nor systematically pleasing, provides the initial steps toward establishing some relationships among concepts. The classification system was again emperically tested and found to account for the data contained in the O.T.U.s. Finally, all of the O.T.U.s were scored independently by two graduate students who had been trained in the use of the scoring definitions and checked by a senior researcher. Differences were discussed and a single score arrived at for each category using a "go-no-go" system.

Bureaucratic Based Classification Schema

A theoretical tri-dimensional taxonomic model was derived by retrieving from the literature those writings, ideas, theories, concepts, and empirical studies that address themselves to an understanding of the behavior of men in organizations, especially those employing the conceptual strategy of *bureaucracy*. The term "bureaucracy" first appeared in social science literature with the writings of *the* Max Weber.

The signal and perhaps most crucial contribution of Weber's theory of bureaucracy was that it provided a conceptual framework for the systematic study and understanding of complex modern organizations. This is to say, the theory accounts for the relationships between organization attributes which permit the observation, classification, and analysis of the social forms that produce these relationships. Thus, at the very core of the Weberian theory are the structural characteristics of bureaucracy and their relationships to each other.

The structural characteristics of a bureaucracy as Weber conceived them might be summarized in outline form:

1. *Specialization*—fixed and jurisdictional areas which are regularly ordered by rules, i.e., by laws or administrative regulations.
2. *Hierarchy*—levels of graded authority that ensure a firmly ordered system of super- and sub-ordination in which the higher offices supervise the lower ones.
3. *Explicit Written Procedures*—all administration based on written documents—a body of officials engaged in handling these documents and files, along with other material apparatus that make up a "bureau" or "office."
4. *Expertise*—all administration by full-time officers who are thoroughly and expertly trained.
5. *Formal Rules*—all administration by general rules which are quite stable and comprehensive.

To Max Weber, the most salient variable in bureaucratic organizations was authority. Accordingly, bureaucratic behavior in an organization can be observed through the "flow of authority." In order that the "flow of authority" become clearly observable, it must be structured. Robert Presthus has presented the structural characteristics of a bureaucracy. They are hierarchy, specialization, oligarchy, cooptation and status. The above characteristics have been subdivided further into their functional aspects (i.e., *hierarchy*—rank, line and staff; *specialization*—office, role expectation; *oligarchy*—power, sanction; *cooptation*—selection, succession; and status prestige, privilege). These provide the structural context within which administrative behavior takes place.

The strategy outlined by Arthur P. Coladarci and Jacob W. Getzels provides the vantage point from which to observe and classify behavior within the structural framework offered by Presthus. They conceive behavior is as taking place in the interactions between individuals and groups in a social structure. For taxonomic purposes the strategy is to note, record, and then classify each observable interpersonal relationship as a distinct specimen or "unit of behavior." Each observation can be operationally viewed as an *administrative dyadic transaction.* By administrative dyadic transaction is meant an interpersonal relationship of two or indeed a plurality of actors in an organization who are engaged in interaction. The interactions that occur within these interpersonal relations can be observed from three dimensional directions: its goal direction, its role parameters, and the nature of the affectivity between the actors. The behaviors, once observed, can be classified as having occurred within an organizational context (i.e., classified horizontally by its structural characteristics and vertically with regard to its three dimensions.)

---

[3] Administrative Performance.

Goal direction dimension (i.e., the purposes or objectives of the organization) is classified as either rational or non-rational. These classes were developed from the Weberian conceptualization of the sources of authority. Similarly, role is viewed as a dichotomy utilizing the Parsonian concepts of functional diffuseness and functional specificity. Finally, the Parsonian concepts of universalism-particularism were used to identify the polarities of the affectivity dimension.

Eight classes of bureaucratic behavior emerged logically from the tridimensional theoretical model, each contained a dichotomous polarity result in six categories. These six can be arranged into eight possible classes which are labeled as follows:

1. RSU—Rational-Specific-Universalistic
2. RSP—Rational-Specific-Particularistic
3. RDU—Rational-Diffuse-Universalistic
4. RDP—Rational-Diffuse-Particularistic
5. NSU—Non-rational-Specific-Universalistic
6. NSP—Non-rational-Specific-Particularistic
7. NDU—Non-rational-Diffuse-Universalistic
8. NDP—Non-rational-Diffuse-Particularistic

The eight classes are distinct and separate, and do not represent a continuum of bureaucratic behavior. Even though the first class, Rational-Specific-Universalistic, may be generally viewed as being highly bureaucratic and the eight class, Nonrational-Diffuse-Particularistic, viewed as being pre-bureaucratic, all eight classes can contain dyadic transactions that are a mixture of bureaucratic and pre-bureaucratic behaviors.

The use of quantile scores enables one to arrange each specimen of behavior along a continuum ranging from a positive pole (bureaucratic) +15 at one extreme, to that of a negative pole —15 (pre-bureaucratic) at the other extreme. The quantile score indicates the degree of bureaucratic behavior in the specimen under inspection. This results in a more refined and discriminating classification of administrative behavior. Another reason for employing quantile ranges is to account for the varying sizes (the number of dyadic transactions) of the O.T.U.s. This enables one to classify and then compare O.T.U.s regardless of the magnitude of behavior present.

Synthesis and Probable Outcomes

As stated in the fifth purpose of the study, it was hoped that it would be possible to synthesize the taxonomies into a single taxonomy of organizational behavior in education. Based on the widely held belief that there is considerable overlap among the theories upon which administration is based, it was assumed that a synthesis would be possible. On the other hand, based on the belief that there are wide gaps in the existing theories, the major problem could have been, that there would be organizational behavior which might overlap but could not be classified because of inadequate theory. Whether or not gaps exist in the the present theories, the classification schema developed from each of the four theories classified all units of behavior. Said another way, each of the four theories was able to classify all the organizational behavior collected in the field study. The gaps were not in the classification generated by the theory. Some may have been gross classifications, yet all data was classified.

Logical and statistical analyses were undertaken in order to test the overlaps which did occur among the taxonomies developed. The detailed procedures are beyond the space limitations imposed here. Briefly theoretical and logical "guesses" were made about overlaps between theories; these were not born out by inspection. Statistically the Phi coefficient was tabulated for all possible pairs of the approximately 450 characteristics scored in the four taxonomies. Only twelve overlaps occurred which were statistically meaningful. Of these twelve the vast majority were schematically predictable. That is to say, that the overlap between the decision-making characteristic of (decision made) could be schematically predicted to overlap with the general systems characteristic (input-decision making). The overlap was not theoretical in nature. One possible theoretical linkage did occur between the bureaucratic taxonomy and the general system taxonomy. All bureaucratic characteristics which account for high bureaucratic performance (high rational, high specific and high universalistic) overlapped with the general systems characteristic (feedback and limited fixed). Theoretically it could be predicted that if an organization were operating in a bureaucratic fashion, feedback would occur in a specific or fixed manner, and would be limited to the kind and process of feedback stipulated by organizational rules.

While this research did not result in a single taxonomy, it has already served as a basis for additional research. Several doctoral dissertations are already underway based on the project, and some proposals for funded projects are being considered. One is now funded for the development of teaching materials from the TOBE project.

Although not useful as a single classification system for abstracting, guides have been presented for classifying within theories. In addition, the study of organizations should be enhanced by this information. Based on the purpose of the investigation, an organizational study can select a theoretical approach and is provided with organizational characteristics through which hypotheses can be generated and data collected. It is the position of the research team that the development of taxonomies in organizational behavior in education has begun; it is in no way assumed that the job is completed.

PROGRAMS FOR TRAINING EDUCATIONAL RESEARCHERS

S. David Farr, State University of New York at Buffalo

Because of my own current concerns I should like to concentrate on the nature of doctoral programs directed to the training of educational researchers. I do not mean to imply that by this concentration that I reject the possibility of training researchers through either pre or post doctoral programs, but it does represent an opinion that doctoral programs will for a considerable period of time be the primary producers of active educational researchers.

Two assumptions underlie my ideas. First, I believe that we must have enough commitment to the training of educational researchers that we are willing to invest in separate programs for educational researchers and for educational practitioners. Second, I am quite willing to assume that the students we admit to our research training programs are intellectually able people, undoubtedly bright enough to complete whatever program we expose them to, and that they have whatever prerequisite knowledge we require for our programs. This allows me to concentrate exclusively on the relationship between the program and the behavior desired in its graduates.

I believe that many current programs are based on two major fallacies. The first is that statistical analysis, measurement, and research design are conceived as the most important part of the training effort; often except for the existence of these the research training program may be no different from the program for practitioners. I would suggest that extreme commitment to, and the exact knowledge of facts and theories of a rather specific applied area is more necessary to the conduct of productive research. The educational researcher must in addition be competent, but not necessarily a specialist, in statistics and design. It would seem to me that an enthusiastic, active, scholarly, research doing professor in the applied field is much more important in the training of a researcher than are the professors who teach him his research skills.

The second major fallacy and perhaps even more important than the first, is an apparent assumption that people who are able to do research *will* do research. It appears that the goals of our program should stress typical behavior, that is, what a person *will* do, while our programs have been designed to make a person *able* to do the tasks of research. While it is true that knowledge and skills are important outcomes of a program, I would suggest that the formation of certain scholarly habits is at least equally important.

There are, then, two types of goals toward which we should direct our doctoral programs, those dealing with knowledge, skills, and understanding, and those specifying the habits which we wish our students to have. I believe that we are currently reasonably successful in meeting the first type of goal, through methods of direct source-student communication (teaching, reading, programmed materials), practice of skills, and professor-student discussion. Our present educational practices also seem to give the student ample opportunity to display knowledge, skill and understanding and to have this reinforced by high marks and the other rewards at the disposal of the university professor.

Scholarly habits on the other hand do not seem to be necessary outcomes of our programs. Where in our programs, for example, do we develop the habit of relying on others for details of research in which one does not feel confident? Where do we develop the habit of high activity in doing empirical research, tempering our ideals of excellence with respect for what can be learned from an imperfect study? Where do we develop the scholarly habit of choosing a specific area and studying it with great intensity for some period of time, to the exclusion of other areas? Where do we develop the habit of sharing research results with immediate peers and with the profession?

I would like to suggest the type of program which might be more successful in developing these habits than most current doctoral programs in education. It would seem to me that we might make some progress with a cognitive approach to this problem, teaching the student that rewards will be given for the types of behaviors just outlined. However, a more direct attack would be to have our programs center around participation in the research process so that the various steps of this process can be attempted and successful performances reinforced. It would seem that a program of participation in research might begin with observation and directed participation in the research of faculty and advanced students. Very shortly the student could be stimulated to make first attempts at various phases of research, perhaps in writing a report of research conducted by another or by preparing a statement of problem or an outline of a research plan. Having made some first attempts in situations simple enough so that success is nearly assured, the student might progress to a stage which could be characterized as guided practice. From there he might move to cooperative research with his professor and then to individual projects or perhaps collaborative work with his peers, in order to transfer some of his dependence from his professor to his peer group.

By engaging the student in such a program, we would have many opportunities to use the rewards at our disposal to reinforce desired behaviors. For example the intensive study of a topic resulting in an excellent review of literature could be reinforced by high marks, distribution of the paper to students and faculty, opportunity to present his findings to students and faculty, personal congratulation and encouragement, etc. We could similarly reward the synthesis of skills and knowledge around a specific problem resulting in a research plan. Other stages of the research process, such as the completion of data collection, the completion of the analysis and the completion of research reports could also be rewarded. This last activity, a very critical one, can often be rewarded for advanced students not only within his university, but by rewards of publication or presentation of papers at state and national meetings.

I would suggest as a final principle that in all of our programs of research participation we take a position of "accentuating the positive" by concentrating on the reinforcement of desired behaviors rather than the criticism of weakness.

What then might be the components of a program which would accomplish the process just outlined? I believe that an important first component would be intensive seminars which would allow a student, especially the advanced student, to concentrate his study for a semester in one or perhaps two large-credit seminars rather than spreading himself over four or five different courses. This would seem to have the advantage of allowing him to display and be rewarded for the intensive study of a particular area, and also the opportunity to follow through on his study by conducting a relevant research project individually or cooperatively with the other students. The opportunity for desired activities to occur and be rewarded seems much greater with this kind of study than in a system where the student must divide his efforts among several more or less related subjects and please several professors.

A second component of the program would be the student's work as a research assistant to an active professor in his major field. Here he may experience the joys and frustrations of the researcher through identification with his professor and by direct rewards given for whatever parts he played in the research process. A third major component of the program would be independent study, individually or in cooperation with a small group of students. This kind of work could emphasize either the intensive study or the research-doing aspect of the program and in either case the rewards of the university may be applied whenever appropriate behavior is observed by the faculty.

In summary, then, I see that two kinds of goals are important in the training of educational researchers. The first is that of knowledge, skill and understanding of his applied area and of research methods. It seems to me that we are currently doing reasonably well in terms of these goals. However, it seems that we have paid little attention in our program planning to goals describing the habits which we hope to develop in our students. I suggest that we have the knowledge to design programs which will develop the desired habits, and should get to work on that task.

TRAINING FOR SCHOOL RESEARCH DIRECTORS

Louis T. Di Lorenzo

The passage of the Elementary and Secondary Education Act has elevated interest in the training of educational researchers to a new high. It has also revived some of the problems and issues long associated with this area; such as, the past and potential contributions of research to the educational practitioner, the relative merit of basic versus applied research, individual as opposed to interdisciplinary team study, the dominance of the psychological base of educational research, and training with emphasis on research methodology or on a discipline. The resolution of these problems is still to be realized; at best, any stand on these issues has been supported by the logic which follows the set of assumptions and values set forth.

The Sixth Annual Symposium of Phi Delta Kappa in 1964 addressed itself to each of these issues. Despite the caliber of the participants, few inroads were made into the problems which could be expressed as either guidelines or recommendations for the soon-to-follow educational research training programs supported under Title IV. However, the past few years have produced studies and reports which have clarified the nature of the training, personnel, and functioning of researchers which can serve to assist those directing training programs. Most significant are the reports by Bargar and Guba, Buswell and McConnell, Clark, Clark and Hopkins, Di Lorenzo, Lazarsfeld and Sieber, Sieber, and Stanley.

These reports and any number of future studies will not settle some of the basic philosophical differences associated with training. They will add only to a refined understanding of the problem and recognition that the positions taken are intrinsically related to the values held. Nevertheless, those training programs under support by Title IV have several clear directives from the legislative intent and language of the Elementary and Secondary Education Act. While the latter is not specific enough for detailed program development, it leaves little doubt that the persons trained in these programs are to function in a capacity to enable the overall success of the programs of ESEA.

THE TRAINING OF EDUCATIONAL RESEARCHERS

Marvin Taylor, Queens College

How does one find and train individuals who have an unquenchable thirst for knowledge, are open-minded, and have an unrelenting drive to subject their ideas to objective scrutiny? I shall attempt to set forth some necessary conditions, that I believe to be worthy of consideration in planning a training program for educational researchers.

When one plans a training program he must take into account the nature of the job the product of the program will ultimately be expected to do. It would be simpler if we could say that the researcher will do research; but what research? As Guba (1964, p. 280) has suggested, research activities should include "development, dissemination, diffusion, and implementation." It seems apparent that to conceive of research in such broad terms implies more than one kind of training program and a multiperson research effort. It seems to me that the multiperson unit would need to include such persons as a design specialist, a subject matter specialist, a writer and interpreter of the basic research, and an administrator to handle funds and coordinate the research unit's efforts in the public schools.

Research activities might be stretched along a continuum of immediately useful knowledge. This would implement innovation to more abstract information which would require extended study in both laboratory and field settings prior to implementation in the classroom. It seems to me that activity in all areas must be pursued if we are to make advances in education.

If research is to provide information useful in creating and evaluating educational innovations, we must prepare individuals who will have the necessary skills and will identify themselves as educational researchers. Let us consider the kinds of individuals we should recruit for training as Educational Researchers. According to our traditions most researchers are not trained as such, but emerge from a field as a result of their doctoral studies. According to Bargar (1964) Educational Administration is the largest category of doctoral majors from which individuals who are listed in the National Register of Educational Researchers are drawn. This is followed by the field of Educational Psychology. Only 1 percent of the respondents indicated that educational research was their major doctoral area. Typically, those individuals who select educational administration, educational psychology, and the other sub specialities are somewhat older and advanced in professional commitment. They have collected more information and have developed more set ways of thinking and processing new information. G. B. Shaw once quipped, "Reading rots the mind." I would suggest that we need to recruit more young people who have just graduated from undergraduate school. They should represent a wide sampling of liberal arts fields and they should have a broad, liberal educational background. We need to attract young people, before they have established large family commitments: 1) because the training program will be long and arduous; 2) because having worked in a field for several years may have reinforced ways of thinking about phenomena which may create problems when one is required to think in more abstract terms; 3) because most social scientists are most productive and creative during their thirties and early forties.

In addition to being bright and young, the college performance of our prospective researcher should be examined for evidence of adventuresomeness, lack of over-conformity, risk taking proclivities and interest in exploring ideas. As an undergraduate our prospective researcher should have demonstrated an awareness of problems, a desire to explore ideas further than the obvious implication, and a stick-to-itiveness in the face of frustration.

Assuming we select individuals with the characteristics mentioned above, the training program must provide experiences which will promote cognitive development and professional attitudes. A portion of these experiences should be in such courses as experimental design, statistics, (at least through multivariate analysis), educational and psychological measurement (for example, questionnaire construction and scaling techniques), philosophy of science, computer technology, curriculum development, educational psychology, philosophy of education, and an academic home base. There should be enough flexibility to permit the student to take course work in mathematics, social science, and other areas of interest which will allow him to expand his ways of thinking.

More crucial than courses, however, are the types of research experiences and the kinds of people the prospective researcher will have the opportunity to identify with and model himself after. As Krathwol (1964,p.92) has suggested,

> Of more importance is the atmosphere where a contagious curiosity prevails; where new ideas are continually fermenting and being challenged; where the new researcher can learn first hand the nature of the compromise in each experimental design; where research is exciting, where careful methods of data handling are modeled; where the student learns how to reduce data to understandable terms; where, in the day of the antiseptic computer—produced data untouched by human hands, he learns to observe carefully and absorb the meaning of his data by talking with his subjects; where he learns the value of studying intensely a few cases and of plotting data, not just noting summary statistics; where he talks to teachers and principals who are actually in the school situation and where he has the freedom to display initiative and is encouraged to delve deeply.

This problem is handled in the large universities through their institutional research offices. With the increase in government grants many individuals have had opportunities to do research in conjunction with experienced and energetic researchers. This, however, is not enough. I feel that institutional research tends to be insular, either at the university or in the public school. It seems to me that there must be a new organization, a consortium, if you will, of universities and public school systems. The resources of both groups must be joined together. The experienced researchers from the universities must work with supervisors and trained personnel from the public schools. The research must be planned and carried out in several districts at the same time. The prospective researcher must work in this setting as an assistant, as an intern and, as a resident researcher. One can conceive of this research unit as similar to the new R & D centers presently being supported by the government. The research unit would be concerned with development, dissemination, innovation, and evaluation.

In summary, let me suggest that recruitment, self-selection, advancement of knowledge, and innovation are all inextricably intermingled. As our research efforts become more exciting, as we approach breakthroughs in our knowledge, as our ideas reach new levels of sophistication, as our techniques of research and our research efforts become of better quality, so too will bright young men and women seek to make educational research a professional area in which they can spend fruitful lives. As we have more to offer we shall be able to be more discriminating in our selection of people. Finally, as better prepared individuals enter the field, and have the opportunity to work in realistic settings, changes will begin to take place in our classrooms and we shall no longer be able to point to the ever expanding, but relatively useless research efforts on the one hand, and the practical but theoretically barren teaching behavior on the other hand. We must begin to pull together or we shall surely drive out the eager minded, young educational researchers of the future.

Bibliography

Bargar, R. R. "Who is the Educational Researcher?" 6th Ann. Phi Delta Kappa Symposium. 1965.

Guba, E. "An Overview of the Symposium." 6th Ann. Phi Delta Kappa Symposium. 1965.

________ & Elam. *eds.* The training and nurture of educational researchers. 6th Ann. Phi Delta Kappa Symposium. 1965.

Krathwohl, D. R. "Current Formal Patterns of Educating Empirically Oriented Researchers and Methodologists." 6th Ann. Phi Delta Kappa Symposium. 1965.

A COOPERATIVE PROGRAM BETWEEN A CITY SCHOOL DISTRICT AND A SUBURBAN SCHOOL DISTRICT

William C. Rock, Herman R. Goldberg, & Joanne E. Lang, City School District of Rochester; L. William Heinrich, West Irondequoit; & James V. Mitchell, University of Rochester

THE PROGRAM

As planned, 25 children in grade one, from William H. Seward School No. 19 with an enrollment of more than seventy percent nonwhite pupils, were sent on a voluntary basis to six neighborhood schools of the West Irondequoit Central School District in September 1965. One of these pupils moved shortly after the school year began.

Free transportation was provided for the children involved by the City School District. Lunchroom facilities and supervision were furnished by the receiving school district. Training sessions were conducted in the receiving schools to prepare teachers for this new experience.

The basic purpose of the program was to improve the educational opportunities for both the children in West Irondequoit and the 25 City School District children. Both groups of children had been attending school in racially imbalanced settings; one predominantly Negro, the other all white. Neither group had had a full opportunity to become acquainted with children from a variety of cultures and, to the extent that this opportunity was lacking, their preparation for life in a democratic society was weakened.

A second major purpose of the program was to demonstrate a metropolitan approach to the solution of the problem of racial imbalance. It is recognized that this is a problem which affects both the city and the suburbs and that there is a need and responsibility for joint action.

This was the first year of a longitudinal program which will include additional children, schools, and grades in later years.

THE SAMPLE

A pool of 58 incoming first grade children from William H. Seward School No. 19 in Rochester were selected as possible participants for the project. William H. Seward School No. 19, located at 465 Seward Street, is an elementary school housing grades K-6. Of the total enrollment of 1100, approximately 78.9 percent of the children are nonwhite. The students selected were average or above average in ability and achievement in the opinion of their kindergarten teachers. Through the use of random sampling 29 of these children were selected to participate in the project. The other 29 children were drawn to serve as a control group. The parents of the children selected were then asked if they desired this opportunity for their children. When a parent of a child in the experimental group said no, a student in the control group was dropped at random. For example, if the parent of the fourth child drawn for the experimental group said no, then the fourth child drawn for the control group was dropped. Since several of the children in the original experimental group were unable to go to West Irondequoit, it was necessary to utilize the control group to secure 25 children to go. In all, six children from the original control group eventually went to West Irondequoit. These were children who had previously been dropped from the control group because the corresponding child in the experimental group had said no to the program. This type of selection did not adversely affect the comparability of the two groups, however, since they were found to have approximately equal results on the Metropolitan Readiness Tests given early in the fall.

ANALYSIS OF PUPIL ACHIEVEMENT

The academic achievement of pupils in the control and experimental groups was compared by utilizing data from three standardized tests administered during the school year. A description of the tests and the subsequent statistical analysis follows.

1. *The Metropolitan Readiness Tests* (Administration Date: October 1965)
   Test 1. Word Meaning (19 items)—Measures pupil understanding of comprehension of language.
   Test 2. Sentences (14 items)—Measures the ability to comprehend phrases and sentences instead of individual words.
   Test 3. Information (14 items)—Measures vocabulary.
   Test 4. Matching (19 items)—Measures visual perception involving recognition of similarities.

a capacity which is required in learning to read.

Test 5. Numbers (24 items)—Measures general number knowledge, including achievement in number vocabulary, counting, ordinal numbers, meaning of fractional parts, recognition of forms, telling time, and the use of numbers in simple problems.

Test 6. Copying (10 items)—Measures visual perception and motor control skills which are required in learning to write.

The total maximum possible raw score for these tests is 100 points.

2. *The Metropolitan Achievement Tests, Primary I Battery, Reading Section Only* (Administration Date: May 2-6, 1966)

   Test 1. Word Knowledge (35 items)—Measures sight vocabulary or word recognition ability.

   Test 2. Word Discrimination (35 items)—Measures the ability to select an orally presented word from among a group of words of similar configuration.

   Test 3. Reading (45 items)—Measures sentence comprehension (13 items) and paragraph comprehension (32 items).

   The scores from each of the three subtests are reported independently.

3. *The Science Research Associates (SRA) Reading Achievement Test, Level 1-2* (Administration Date: May 9-13, 1966)

   Test 1. Verbal-Pictorial Association (48 items)—Measures the ability to understand words, phrases, and sentences and to differentiate between words that look alike.

   Test 2. Language Perception (125 items)—Measures the ability to discriminate between pairs of words having similar sounds, to identify pairs of identical words, and to associate spoken with written words.

   Test 3. Reading Comprehension (36 items)—Measures the ability to understand a central theme and main idea, draw logical inferences and grasp minor details from varied reading selections.

   Test 4. Reading Vocabulary (41 items)—Measures the ability to understand the meaning of words in context. The vocabulary test uses the same reading passages as the comprehension subtest.

The scores from each of these four subtests are reported independently.

The statistical technique used to analyze these data was a one-way analysis of covariance with the total score from the Metropolitan Readiness Tests as the covariable and each subtest of the two reading achievement tests (Metropolitan and SRA) as an independent criterion measure. In all, seven analyses of covariance were run.

The results of these analyses are summarized in Table I on page 60. The mean raw scores have been converted to grade equivalents so that practical as well as statistical significance may be studied. Statistical significance was found for three of the seven analyses (SRA Language Perception, Reading Comprehension, and Reading Vocabulary). In all three instances, the differences in the adjusted means favored the experimental group.

In summary, it may be concluded that the reading achievement of the pupils transferred to West Irondequoit, as measured by the Metropolitan Achievement Tests and the SRA Reading Tests, was not adversely affected by the transfer. In three of the seven instances, it was significantly better than the reading achievement of the control pupils who remained at William H. Seward School No. 19.

ANALYSIS OF SOCIOMETRIC DATA

After the transferred pupils had been in West Irondequoit for two months, a series of sociometric questions was administered to all pupils in the eight classrooms involved in the program. A standard procedure was used in administering the questions. The administrator would first spend some time observing the class, becoming familiar with the seating arrangements and the names of the pupils. A seating chart was then developed for recording the responses to the sociometric questions. The questions were asked at a time when the pupils were doing seat work. The administrator called the pupils up individually to a corner in the front of the room and asked each question orally. Because the other pupils were doing seat work, with heads averted, the pupil being questioned had a visual reminder of the pupils in the room and yet had privacy in giving his responses.

The questions were administered during the months of November, December, and early January and a second time during the month of May. No strong effort was made to restrict the number of nominations a pupil made in response to a question but after three or four nominations, the pupil was encouraged to go on to the next question. There was absolutely no reference made to race or skin color in any of the questions.

The nominations were tabulated in the form of a sociometric matrix. The observed frequency of nomination was then computed for nonwhite and white pupils separately and summated across all eight classes. Chi square analyses were then conducted to compare these observed frequencies for the two groups with the expected frequencies based upon the proportion of nonwhites and whites constituting the total group of pupils in the eight classes.

TABLE I

SUMMARY OF ANALYSIS OF COVARIANCE DATA MEAN RAW SCORES AND GRADE EQUIVALENTS[1]

Criterion Tests		**Group**			
		Experimental		**Control**	
		R.S.	**G.E.**	**R.S.**	**G.E.**
SRA Reading 1-2	Unadj.	14.0	1.7	12.8	1.6
Verbal-Pictorial Association	Adj.	14.3	1.7	12.5	1.6
SRA Reading 1-2	Unadj.	86.1	1.7	70.7	1.3
Language Perception	Adj.[2]	86.9	1.7	69.8	1.2
SRA Reading 1-2	Unadj.	14.0	1.7	10.4	1.2
Reading Comprehension	Adj.[2]	14.2	1.7	10.2	1.2
SRA Reading 1-2	Unadj.	12.6	1.9	9.3	1.4
Reading Vocabulary	Adj.[3]	12.8	1.9	9.0	1.4
Covariable					
Metropolitan Readiness Tests		72.6		74.1	
Number of Pupils		24		19	

Criterion Tests		**Group**			
		Experimental		**Control**	
		R.S.	**G.E.**	**R.S.**	**G.E.**
Metropolitan Achievement Tests	Unadj.	22.2	1.7	22.6	1.8
Word Knowledge	Adj.	22.3	1.7	22.5	1.8
Metropolitan Achievement Tests	Unadj.	23.0	1.8	23.4	1.8
Word Discrimination	Adj.	23.2	1.8	23.2	1.8
Metropolitan Achievement Tests	Unadj.	22.6	1.9	22.5	1.9
Reading	Adj.	22.7	1.9	22.4	1.8
Covariable					
Metropolitan Readiness Tests		73.5		74.1	
Number of Pupils		22		19	

[1] Tested at grade placement 1.8.

[2] Difference between experimental and control groups significant at .01.

[3] Difference between experimental and control groups significant at .05.

The questions asked and the results of the chi square analyses are summarized in Table II. Significant differences were found for Question 1 on the initial testing (fewer nonwhites were nominated as best friends than would be expected) but not on the follow-up spring testing. For Questions 4 and 6, significant differences were found on both administrations (more nonwhites were nominated for getting into fights and not wanted as best friends than would be expected). On these two particular questions, there were three classes where the nonwhite nominations were large enough to make the difference significant for the total group. In the remaining five classes, differences were slight. Regarding Question 5, significant differences were found on the spring testing (more nonwhites were nominated for being silly than would be expected) but not on the initial testing. It is interesting to note that for Questions 2, 4, 5 and 6, there were several white pupils who received just as many nominations, and in some cases even more, as did the nonwhite pupils.

On the basis of these data, it is not possible to judge whether the results achieved here, particularly for Questions 4 and 6, are due to race or to the status of the nonwhite pupils as newcomers to West Irondequoit. In some instances, these differences may even be due to the sex of the pupils (e.g. boys getting into fights). In general, the Rochester children seem to be making an excellent adjustment. It is planned to continue gathering sociometric data on transferred pupils as they progress through school.

TABLE II

SUMMARY OF CHI SQUARE ANALYSES FOR SOCIOMETRIC QUESTIONNAIRE

Question	Date of Administration	Frequency				
		Nonwhite		White		
		Obs.	Exp.	Obs.	Exp.	χ^2
1. Which children in the class are your best friends?	Fall	41	71.0	623	593.0	13.75[1]
	Spring	76	89.2	758	744.8	2.18
2. Which children in the class always seem unhappy and sad?	Fall	19	20.2	170	168.8	.03
	Spring	32	25.1	203	209.9	2.13
3. Which children in the class would you *like* to have as best friends?	Fall	47	48.5	406	404.5	.02
	Spring	62	62.1	518	517.9	.00
4. Which children in the class are always getting into fights?	Fall	60	38.2	297	318.8	13.31[1]
	Spring	79	47.6	366	397.4	23.19[1]
5. Which children in the class are always being silly?	Fall	43	39.3	324	327.7	.30
	Spring	64	48.1	386	401.9	5.89[2]
6. Which children in the class would you *not* want as best friends?	Fall[3]	59	24.5	160	194.5	52.43[1]
	Spring	69	45.9	360	383.1	13.02[1]

[1] Difference between observed and expected frequencies significant at .001.

[2] Difference between observed and expected frequencies significant at .05.

[3] Administered in six of the eight classes.

Classroom Incidents, Interracial Experiences

The Intercultural Enrichment Program is designed to provide opportunities for intercultural experiences, as a means for improving interracial and intercultural relations. Incidents reported by teachers offer examples of the kinds of opportunities which occur at the first grade level.

1. One of the transferred pupils reported another child had said a "bad word" to him. In talking with him, the teacher found out the word was *Negro*. The teacher suggested that he discuss this with his father, and if his father agreed that it was a "bad word," the teacher would talk to the child who said it. The next day the boy reported to the teacher that his father had explained it was not bad, and that it was all right for people to say.
2. A white girl sitting next to a Negro boy at the end of a tiring day noticed him sucking his thumb. She leaned over and quietly inquired as to the flavor—"chocolate or vanilla?"—then popped her own thumb in her mouth to check on its flavor.
3. In referring to a Negro boy's actions on the playground, a white boy referred to him as "that brown kid." The term was used descriptively, not as a derogatory remark.
4. One of the Negro boys was fascinated with the silky hair of the girl who sits next to him. The boys in the same classroom wanted to touch the Negro boy's close-cropped head.
5. In choosing another to accompany him as messenger, a Negro boy selected another Negro boy. A white boy commented, "Why does he always choose a Negro?" The teacher felt the boy was disappointed, and blamed skin color.
6. In discussing differences in families in a receiving school classroom, children included color of eyes and hair and nationality. Skin color was not listed as a difference by the children.
7. The making of Indian headbands was a project in one classroom. A white child suggested that the three Negroes in the classroom would especially enjoy this, as they were Negroes, and dark-skinned, as are Indians.
8. A principal observed a lesson on differences, in which the teacher used books as examples and then switched to pupils, asking two boys to stand up. Group responses noted differences in colors: shoes, socks, pants and shirts. The teacher was

after differences in size and height. The fact that one of the boys was Negro and the other white did not enter the discussion.

Other examples mentioned play activities, a Negro teacher on educational television, and a visit to one of the schools by a predominantly Negro first grade class from a city school, as intercultural experiences.

Social Adjustment of Pupils

As indicated, many of the staff members commented on the way in which Negroes and whites "got along." Most of the observations were general; "children are making friends with one another," "the class plays together with no apparent prejudice," "the Negro children are participating very effectively," or comments of a similar nature. Some comments listed specific friendships.

For two of the eight classrooms, some comments were made about the Negro children "sticking together;" in at least one instance, the teacher felt that this behavior limited the effectiveness of the program.

In most instances, the comments described social adjustment as normal, or noted the interaction between Negroes and whites.

Conclusion

The information provided appears to reflect a rather "normal" situation for most of the pupils involved in the educational program. Expected individual differences are apparent.

The need for changes in administrative procedures, and attention to other phases of the Intercultural Enrichment Program, such as avenues for community involvement and expansion of opportunities for Negro pupils to participate in out-of-school activities, is apparent.

An expansion of opportunities for wholesome interracial experiences is evident, particularly in the comments of teachers concerning the face-to-face relationships among Negro pupils and white pupils. The opportunities for teacher-directed learning experiences in connection with intercultural enrichment were limited in number, as expected at the first grade level. While some of the potentially negative incidents are in need of continued study, they do not suggest the need for significant changes in the overall program.

As a means of evaluation, the distribution and collection of this observation form provided a valuable resource for those responsible for keeping in touch with and reporting on the progress of the program. It has also provided information which will be used in planning for future years, and it helps to provide information on the effectiveness of the program.

It must be recognized that this form does structure the responses of staff members; it also suggests the need for looking at the Negro pupils as Negroes. From this standpoint, and from the standpoint of the time required of staff members to complete it, restricted use of the form is suggested, and it probably should be eliminated as soon as possible.

SUMMARY

It is important to note that this is an interim report of a longitudinal study and that no major conclusions can or should be made at this time. The study will continue next year with a new group of pupils entering grade one and the present group advancing to grade two. The progress of these pupils and subsequent groups will be carefully evaluated.

However, it is possible to conclude on the basis of data collected and analyzed to date that the program is working well, and that children involved at this time are benefiting from the experience.

THE ELEMENTARY AND SECONDARY EDUCATION ACT OF 1965—TITLE I

Irving Ratchick, New York State Education Department

On April 11, 1965, Congress enacted Public Law 89-10 known as the Elementary and Secondary Education Act of 1965 and on September 23, 1965 the President of the United States signed the Appropriation Bill. This marked the launching of a great effort to break the cycle of poverty through education.

Title I, the major thrust of the Act, has as its Declaration of Policy the following:

> In recognition of the special education needs of children of low-income families and the impact that concentrations of low-income families have on the ability of local educational agencies to support adequate educational programs, the Congress hereby declares it to be the policy of the United States to provide financial assistance (as set forth in this title) to local educational agencies serving areas with concentrations of children from low-income families to expand and improve their educational programs by various means (including preschool programs) which contribute particularly to meeting the special educational needs of educationally deprived children.

The allocation for New York State for the first fiscal year of 1966 was approximately $110,000,000. The allocation was based on two criteria: (1) the number of children ages 5-17 who reside in families where the income was $2,000 or less per year, as indicated by the Bureau of Census data of 1960, and (2) the number of children ages 5-17 who reside in families where the payments under the program of Aid to Families with Dependent Children (AFDC), Title IV of the Social Security Act, was $2,000 or more as determined by the 1962 AFDC data. These two figures added together times one-half the average pupil expenditure for education of the previous year, which in New York State amounted to approximately $365, gave the amount of the allocation. It is interesting to note that of the $110,000,000, approximately $75,000,000 is given over to the Big Six (New York, Buffalo, Rochester, Syracuse, Albany, and Yonkers).

Congress was very definite that this Act was intended to provide special programs and services for children who were educationally disadvantaged, primarily from pockets of low income. The Act was not intended to provide general aid to schools. Moreover, significant parts of the Act consider the involvement and planning with nonpublic schools in order that disadvantaged children who attend nonpublic schools should receive the benefits as well. Furthermore, it is mandated that there be planning and cooperation with the community action groups sponsored under the Economic Opportunity Act, since there are like projects fundable under both Community Action Programs and Title I of ESEA.

School districts known as local educational agencies send in projects and designated pockets of poverty within their areas. Where a school district because of its geographical area was heterogeneous in its makeup, the entire school district could be considered as a single unit. A significant point of operation for this year is that funds which were not used by one local educational agency could not be transferred to other local educational agencies. An amendment has been submitted to Congress, and if it is passed, next year we shall be able to transfer unused funds where needed.

Although the basis for allocations was economic deprivation determined by 1960 and 1962 data, programs and services had to be provided on the basis of educational disadvantagement.

Educationally deprived children are designated by the Office of Education thusly:

> Educationally deprived children are children whose educational achievement is below that normally expected of children of their age and grade, including children who are handicapped because of physical, mental, or emotional impairment.

Once the allocation had been determined, services and programs were to be given for the educationally disadvantaged. But, how do you determine if a pupil is disadvantaged? Is it by data from standardized tests, teachers' observations and recommendations, appearance, parental occupation, or is it by a combination of any of these or other factors?

In order to assist school districts determining who were the educationally disadvantaged, I prepared a paper entitled, *Identification of the Educationally Disadvantaged.* In the publication, criteria for identification of the educationally disadvantaged were placed under two general headings: (1) quantitative and (2) qualitative. Quantitative data would include such items as the results of mental ability tests, achievement test data, including scores in reading and other tool subjects and in statewide tests, school grade and age comparisons, pupil-personnel services information and health status and handicaps. Qualitative data which may be more closely related to the intent of the Act would include social and cultural factors, teacher observations and judgments, other staff and community agency evaluations, results of surveys such as that of parents' occupations, employment status, attitudes, education, and additional findings of research pertaining to cultural and educational disadvantagement.

In order to establish baseline data as required by the Act, the New York State Pupil Evaluation Program was initiated this past year. All students in grades 1, 3, 6 and 9 in public and non-public schools were given the same tests. There were reading and number readiness tests for grade one, elementary school reading and arithmetic tests for grade three and six, and minimum competence reading and arithmetic tests for grade nine. This program will be repeated each year and testing is now proceeding for the current year. Priorities for the projects were given as follows:

1. Direct immediate benefits to pupils. This included activities such as tutoring, classes in remedial reading, mathematics, cultural enrichment activities, additional personnel, and so forth.
2. Programs for professional development. This included inservice education, workshops, and so forth. If real progress is going to be made in working with the disadvantaged, the classroom teacher who has most daily contact with pupils should be given the opportunity to acquire deeper understanding and knowledge of the nature and needs of the educationally disadvantaged.
3. Curriculum materials and supplies to provide greater differentiation of instruction.
4. Equipment.
5. Construction which was practically nil.

We had and are having a most active year in FY 1966 ending June 30, 1966, 1,461 projects were received by the Department involving an expenditure of approximately $101,400,704.12 out of an allocation of $110,000,000. Eight hundred twenty-nine local educational agencies participated in the program which represents 93% of the public school districts of the State. In addition, $2,700,000 was made available to State-operated and State-supported schools for the handicapped under an amendment (P.L. 89-313) to Title I, ESEA. These schools included the 14 schools for the blind and deaf under the supervision of the State Education Department as well as schools under jurisdiction of the Division of Youth, Department of Health, and the Department of Mental Hygiene. For example, the schools at Batavia, New York, Institute for the Blind, and LaSalle School received allocations. In 1967, as of September 14, 1966, 413 project applications have been received amounting to $22,211,126.66. An amendment present in Congress will provide allocations for institutions for neglected and delinquent children. We have approximately 117 institutions which fall in the category. Approval of the amendment will involve the Department of Correction and Social Welfare in addition to the other State agencies.

Both houses of Congress are presently considering amendments to the Act. Each section is recommending a much greater amount than what was in effect last year. The Senate, moreover in its bill, is recommending a new title, Title VI, which will be concerned specifically with handicapped children. As is well known, there are many pieces of federal legislation which are concerned with the handicapped and the purpose of this section will be to try to coordinate many programs for the handicapped. It is being recommended under Title VI of the Senate Bill that a Bureau for Education and Training of the Handicapped be established within the Office of Education.

There is no limit to the creativity which can be exercised by school districts. Projects submitted by school districts must represent an initiation of a new project, or extension, or improvement of existing projects. Some possibilities for projects under Title I include the following:

> Inservice training for teachers and other professional staff presently employed in schools serving the disadvantaged children—institute, workshops, etc.; supplementary employment of staff members with special skills and additional teaching personnel to reduce class size and supplement instruction; programs to recruit and train teachers, psychologists, attendance teachers, physicians, school nurse-teachers, social workers, home and school visitors, reading specialists, speech therapists, guidance counselors, school-job coordinators, and teacher-aides for schools serving disadvantaged students.

Employment of personnel needed to achieve coordination among the local educational agencies concerned with the welfare of children are possible.

Additional activities may include establishing: programs of parent and home visits; counseling services for parents for preschool and handicapped children; programs for early identification of pupil aptitudes, abilities, disabilities, and development of curriculum and teaching procedures; resource rooms and itinerant teacher programs; bookmobiles, libraries, and facilities for the development of curriculum materials for disadvantaged children; supplemental health and food services; mobile units and itinerant specialist service for special education classes in sparsely populated areas; preschool classes including those for deaf, blind or multiple handicapped; physical fitness activities to improve mental and physical health; summer programs including remedial, cultural enrichment, and recreation; provision for the development of the talents of gifted children among the disadvantaged; English programs for non-English-speaking children from deprived homes; remedial programs especially in reading, mathematics and tutoring; special classes for physically handicapped, mentally handicapped, emotionally disturbed and socially maladjusted children; after-school study centers; provision for development and purchase of communication media and other resource materials; program to assist potential dropouts including tutoring and work-study for unemployed, out-of-school youth, between the ages of 16 and 21; workstudy programs, on-the-job training, and part-time employment with stipends; provision of a staff member to supervise vocational training programs for handicapped youth which provide part-time job placement; on-the-job supervision supplemented by special class training with emphasis on tutoring in different areas discovered on the job.

A unique program instituted by the Rochester City school district was a project to increase intercultural understanding through pupil transfer plans. The most important objects of the program were to reduce racial imbalance through voluntary transfer plans, to increase intercultural understanding, and to improve the achievement, attitude and self-concept of the children involved. Under this program, children in schools which are racially imbalanced were permitted to transfer voluntarily to receiving schools in other parts of the city and in the suburbs as well. Free transportation was provided for the children. Approximately 900 children were served under this project; the total estimated budget was $108,492.

A second project worthy of special notice, in the city of Niagara Falls, was devoted to ornamental horticulture providing occupational skills and training for landscaping, floriculture, floral design and arrangement, and nursery work. The curriculum included instruction in the classroom and in the field. Some 40 nonpublic school children from poverty areas participated in the project in addition to 200 public school enrollees and 50 children not enrolled in any school. The total estimated budget was $54,955.

In a program in New York City involving mothers of educationally deprived children as volunteers, a Teacher-Moms program was instituted to provide an educational and supportive setting during the summer vacation period for seriously emotionally ill young children. A minimum of professional staff was employed. The Moms program was designed to help those who have experienced difficulties in social relationships and who manifested such learning problems that a regular group situation was ineffective. A one-to-one learning experience was set up instead with what is known as teacher-mom. Each child had two volunteer teacher moms who alternated two mornings a week. It was with these teacher-moms that a one-to-one learning situation is able to exist. Academic and social activities are encompassed within a broad range of objectives for each child as he works with his teacher-mom. Opportunity for both structure and creativity was provided. Appropriately, the program for the child was tailormade to his emotional needs and learning readiness.

This year, dissemination of desirable practices and procedures, evaluation, and followup activities by the staff of the Department will be stressed.

The following represent a few of the highlights:

1. Education of the disadvantaged can be planned on a continuum beginning with preschool education, extending through elementary and secondary schools, and with consideration to post high school education. This has a tremendous potential for teachers to learn about all children.
2. Education in each school district became more cohesive in that public and nonpublic schools were involved in planning of programs and services for disadvantaged and handicapped children. This closeness gained an impetus from the New York State Pupil Evaluation Program cited previously. Immediate benefits were given to pupils both in public and nonpublic schools.
3. Staff members became more cognizant of learning characteristics and needs of all pupils as they studied the nature and needs of the educationally disadvantaged. Professional development of staff members was accelerated through opportunities for school districts to provide in-service education programs. In focusing upon the educationally disadvantaged, the staff was able to acquire a

deeper understanding of the community including socioeconomic and cultural factors. The staff was brought into closer contact with parents.

4. Supplemental special programs and services were made available to physically, mentally and emotionally handicapped children.
5. Schools were able to either develop or acquire desirable curriculum materials and improve supportive services to assist the educationally disadvantaged with their learning experiences. New insights in teaching the disadvantaged were acquired. Integration activities were fostered.
6. Broader horizons for the disadvantaged were made possible by cultural enrichment activities such as visits to museums and attendance at concerts and dramatic performances, including the Lincoln Center for Performing Arts and the Saratoga Performing Arts Center. A project submitted includes the use of television receivers and other educational communication media in public and nonpublic schools having a high concentration of disadvantaged pupils.
7. The acquisition of desirable attitudes and skills for the world of work has received a greater impetus. One project submitted provides work assignments for Neighborhood Youth Corps enrollees for about four hours a day and an educational program for about two hours. Another project submitted is designed to assist socially disadvantaged and academically underachieving pupils in junior high schools consider the development of saleable prevocational skills, and career guidance, in special classes.
8. Increased emphasis was given to identification and development of talent among the disadvantaged for post-high school education. The College Discovery and Development Program initiated by the Division of Teacher Education of City University and the city school system has about 580 ninth grade disadvantaged pupils participating in the project. Through special programs of instruction, guidance and cultural enrichment, students who completed the high school program are guaranteed admission to one of the units of the City University. College students receiving compensation under the College Work Study Program assist in tutoring these pupils.
9. Community resources, higher education institutions, and Federal, State, and local agencies are being utilized more effectively. A project submitted by the City of New York, entitled the School—University Teacher Education Center (SUTEC) involves both the schools and Queens College of the City University. The primary objective of the project is the comprehensive training of teachers for disadvantaged children.

Title I of ESEA and its amendments have provided an increased impetus for services and programs for disadvantaged children. Much has been done and much more can be done. We need special education programs for those who do not have any, and we need improved programs for those who do have them. The maximum utilization of community resources and Federal, State and local agencies can help usher in a new era for the educationally disadvantaged. Education is in the vanguard; it has a great challenge and a greater responsibility. I am positive that each succeeding year will see new advances and benefits for children, youth, and adults.

Research Reports

THE FACILITATION OF PROBLEM SOLVING AND VERBAL CREATIVITY BY EXPOSURE TO PROGRAMMED INSTRUCTION

Richard E. Ripple, Cornell University & John S. Dacey, Boston College

A verbal creativity battery and a behavioral version of the two-string problem were administered to 50 eighth-grade *S*s who had been randomly assigned to an instructional or noninstructional treatment. The instructional treatment consisted of a set of ten programmed lessons designed to facilitate verbal creativity and problem-solving behavior. In addition, 36 instructional-treatment *Ss* and 50 noninstructional-treatment *S*s received a paper and pencil pictorial version of the two-string problem. Instructional-treatment *S*s receiving the behavioral version of the two-string problem solved it significantly faster than noninstructional-treatment *S*s. Also, instructional-treatment *S*s scored significantly higher than noninstructional-treatment *S*s on three of the tests in the verbal creativity battery. *S*s receiving the behavioral version of the two-string problem were superior to *S*s receiving the pictorial version in percent of and time to solution. Results are interpreted as supporting nonspecific transfer effects of the direct training of generalized problem-solving skills through programmed instruction.

RELATIONSHIPS OF SOME STUDENT CHARACTERISTICS AND VERBAL CREATIVITY

John S. Dacey, Boston College & Richard E. Ripple, Cornell University

In this exploratory study, the relationships of 10 student characteristics and five measures of verbal creative performance of 200 eighth graders were examined. Student characteristics were: sex, mental age, sex-role identification, anxiety, constructive and unconstructive compulsivity, independence, dogmatism, convergent thinking, and organizational climate of the school. Measures of verbal creativity were of the paper-and-pencil type. They were: imaginative story writing, flexibility, originality, fluency, and standardized total score on these four.

Correlations of each of the student characteristics with each of the measures of verbal creativity were computed. In addition, stepwise multiple correlation equations were computed for each measure of verbal creativity, using the combined student characteristics as predictor variables.

With these treatments of the data, it was possible to determine the strength and direction of each of the relationships, as well as the relative importance of each of the student characteristics in the prediction of each of the measures of verbal creativity. Major findings were:

1. Unconstructive compulsivity (as was expected) is negatively correlated with verbal creativity, but constructive compulsivity has a positive correlation.

2. An "open" school organizational climate is negatively correlated with verbal creativity.

3. None of the student characteristics examined have relationships with verbal creativity as high as those reported in the literature subjects, older and younger than the adolescents in this study. Discussion of the results and their implications is presented in terms of adolescent development.

THE RELATIONSHIP OF ANXIETY, CREATIVITY, AND INTELLIGENCE TO SUCCESS IN LEARNING FROM PROGRAMED INSTRUCTION[1]

Robert P. O'Reilly, Richard E. Ripple, & Donald J. Treffinger, Cornell University

The objectives of the present study were to estimate the relative importance of anxiety, verbal creativity, and verbal intelligence to achievement in a linear, constructed-response style program. Based on an analysis relating certain characteristics of the programed learning task to these learner characteristics, the following predictions were made: (1) anxiety would contribute nonsignificantly to achievement with programed materials; (2) verbal creativity would contribute negatively to achievement with programed materials; (3) intelligence would contribute positively to achievement with programed materials, but would not be a major factor as it is under conventional modes of instruction.

In addition to the above objectives, the design and analysis of the research made it possible to secure evidence regarding the following:

1. To what extent does pre-instructional knowledge of the learning material taught in the program contribute to the *Ss*' final achievement?
2. To what extent do the double interactions of anxiety, creativity, and intelligence contribute to achievement with programed materials?

Subjects

Subjects for the research consisted of 80 boys and 85 girls in 9 sixth grade classes from two school systems in upstate New York.

Procedures

The materials used for instruction consisted of a published program entitled, *Latitude and Longitude* (Haring, Haring, and Drum, 1963); a ten-page mimeographed response booklet; and a set of instructions designed to standardize use of the program across classrooms and minimize teacher-student interaction during instruction. The program consists of ten lessons presented in a linear style, and requires constructed responses in each frame. For ten consecutive school days, *Ss* received instruction from the program during a one-half hour class period set aside for this purpose.

Prior to the initiation of instruction, *Ss* were given the following tests:

1. The Test Anxiety Scale for Children (TASC).
2. The Lie Scale for Children (LSC).
3. A verbal creativity battery composed of four subtests named: imagination, flexibility, originality, and fluency (Dacey and Ripple, 1965).
4. The Lorge-Thorndike IQ Test (Level-III, Form A).
5. A criterion test specially constructed for the research to measure knowledge of the learning material taught in the *Latitude and Longitude* program.

Subjects were readministered the TASC, the LSC, the creativity battery, and the criterion test at the conclusion of instruction. The *Ss*' LSC scores were adjusted to have the same variance as their TASC scores, and the two scores were combined to obtain the anxiety score used in the study. This score was designated the $\text{TASC}_{\text{adj.}}$ (adjusted) score.

Reliabilities of the test scores for the total sample were: (1) .91 (split-half) and .75 (two-week, test-retest) for the criterion test; (2) .66 (two-week, test-retest) for $\text{TASC}_{\text{adj.}}$ scores; and (3) .39 for imagination, .55 for flexibility, .59 for originality, and .65 for fluency (all two-week, test-retest). Intrascorer stability coefficients for the creativity tests ranged from .68 to .97.

Analysis and Results

The *Ss* raw scores on the measures of anxiety, creativity (four test score), intelligence, pretest criterion, and posttest criterion scores were converted to z scores. Sex of S (expressed as a z score) was included as a control variable. Nine interaction scores were calculated for each S from all the possible cross products of the *Ss*' z scores on the $\text{TASC}_{\text{adj.}}$, the four creativity subtests, and the IQ test. A stepwise multiple regression analysis was then carried out to assess the relative contributions of the eight independent linear variables and the nine interactions to the prediction of

---

[1] This research was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education, and Welfare.

the criterion. A significance test (McNemar, 1962) was applied to the regression coefficients obtained in each step of the regression analysis to determine if any of the independent variables had anything unique to contribute to the prediction of the criterion.

Pretest and posttest means on the criterion test were respectively 39.81 and 57.16 percent of the total possible score. A *t* test for correlated means indicated that the mean gain in achievement (approximately 18 percent) was highly significant ($t = 16.58$).

Seventeen regression equations were obtained in the stepwise regression analysis, indicating that all of the independent variables were combined in the multiple regression equation. The independent variables which contributed significantly to the prediction of achievement, when combined in the multiple regression equation were: pretest criterion ($p < .002$), verbal IQ ($p < .002$), anxiety ($p < .02$), and the interaction of originality and IQ ($p < .02$). The relative importance of these variables may be ascertained from their beta weights obtained in the four-variable regression equation, given below:

.431 pretest criterion
.436 verbal IQ
—.125 anxiety
—.106 originality × IQ

These data show that verbal IQ and pretest criterion were of approximately equal importance as contributors to the prediction of the criterion. The weight obtained for anxiety is negative, and considerably smaller, relative to the weights obtained for pretest criterion and verbal IQ. The interaction of originality and IQ functioned in the regression equation as a suppressor variable, as indicated by its negative weight in the regression equation, and the fact that it correlated at a near zero level with the criterion variable (cf. McNemar, 1962).

The total criterion variance accounted for by the combined predictors, pretest criterion, IQ, anxiety, and the interaction of originality and IQ was approximately 70 percent. Pretest criterion and IQ each accounted for approximately 32 percent of achievement variance. The remaining portions of the criterion variance were attributable to anxiety (approximately 6.6 percent) and error of prediction (approximately 30 percent).

Discussion

It is evident that expectations concerning the contributions of anxiety, creativity, and intelligence to achievement from the program used in the research were not supported by the data. The relatively large contribution of IQ and pretest achievement to final performance indicates that the program was not equally effective for students differing on these characteristics. These findings, in addition to similar findings reported by others (Gagne and Paradise, 1961; Gotkin, 1963; Moore, Smith, and Teevan, 1965), suggest the need for more flexible use of programed materials in the typical heterogeneous classroom.

The negative contribution of anxiety to achievement with the program is consistent with previous studies of anxiety and school achievement (Feldhusen and Klausmeier, 1962; Ruebush, 1963; McCandless and Castaneda, 1956). However, it appears that this result was largely a function of the difficulty of the criterion measure as indicated by the nearly comparable correlations of $\text{TASC}_{\text{adj.}}$ scores with pretest ($r = -.52$) and posttest ($r = -.53$) criterion scores. Thus, the negative contribution of anxiety to achievement in the present study does not necessarily reflect on the relative appropriateness of the programed method of instruction for the more anxious student.

Bibliography

Dacey, J. S. & Ripple, R. E. The Development of the Students Abilities Survey (Verbal Creativity Battery). Unpublished manuscript, Cornell Univer., 1965. 30 pp.

Feldhusen, J. F. & Klausmeier, H. J. "Anxiety, Intelligence, and Achievement in Children of Low, Average, and High Intelligence." Child Development. 33: 403-409. 1962.

Gagne, R. M. & Paradise, N. E. "Abilities and Learning Sets in Knowledge Acquisition." Psychol. Monogr. 75: 518: 14. 1961.

Gotkin, L. G. "Individual Differences, Boredom, and Styles of Programing." Programed Instruction. 3: 1: 11. 1963.

Haring, L., Haring, Darlene, & Drum, J. Latitude and Longitude. Coronet Instructional Films. Chicago. 1963.

McCandless, B. R. & Castaneda, A. "Anxiety in Children, School Achievement, and Intelligence." Child Development. 27: 379-382. 1956.

McNemar, Q. "Psychological Statistics." John Wiley. New York. 1962.

Moore, J. W., Smith, W. I. & Teevan, R. I. "Motivational Variables in Programed Learning; the role of need achievement, fear of failure, and student estimate of achievement as a function of program difficulty." Unpublished report, Cooperative Research

Grant No. 7-48-0070-149.1, Office of Education, U.S. Department of Health, Education, and Welfare, 1965. 76 pp.

Ruebush, B. K. Anxiety in Child Psychology, the Sixty-second Yearbook of the NSSE. Univ. of Chicago Press. Chicago. 465-516. 1963.

1,000 PROGRAMS OF SELF-INSTRUCTION FOR ALL AGES: AN ON-GOING NATIONAL SURVEY[1]

Jonathan W. Varty, Brooklyn College, City University of New York

Purpose[2]

The purpose of this survey is to locate programs of self-instruction for all ages:

1) to present an up-to-date resource listing of published or manufactured programs which are for sale to the public on a nonrestricted basis
2) to present pertinent descriptive and factual data for each program
3) to attempt the formulation of objective evaluating criteria

History[3]

This on-going research is a refinement and an extension of two previous studies reported at ERANYS convocations; the first in 1963 and the second, 1964.

Method[4]

The mailer consisted of a one page data sheet, a one page covering letter, and a preprint: *Programs of Instruction for Adults.* Data were requested from 200 sources: author, editor, publisher, manufacturer, businessman, industrialist, research and professional organizations, and the United States Government.

Limitations

1) The addressees were limited to the U.S.A. and Canada.
2) Unclaimed mailers numbered 46; it was assumed that 154 mailers reached their destination; 84 or 55 percent of the mailers were completed and returned.
3) The cut-off date on this report: November 1, 1966; follow-up letters are being mailed.
4) "In-house," custom, and Armed Forces programs, not for sale or available to the public on a non-restrictive basis, were excluded.

Nature and Scope of the Data

In addition to the author, title, date, and publisher or manufacturer, this survey will include data in the following categories: *format* i.e. textbook, film, tape, et cetra; *system* i.e. linear, branched et cetra; *mechanics* i.e. number of pages, frames, price, et cetra; *machine* i.e. optional or required; *auxiliary* i.e. index, tests, et cetra; *number of revisions* prior to publication; availability of a *teacher's manual;* nature of *target population,* prerequisites, if any; *number of study hours* to complete; and the number and level of *persons field-tested.*

Findings

1) *Age or school level break down*

The 936 programs as detailed in Table I offers the following break down for age or school level; senior high school 23 percent, junior high school, 21 percent. forty-four percent of the programs were designed for secondary schools.

Programs designed for college students constitute 21 percent, those planned for graduate students and other adults make up 16 percent or 37 percent of the programs are designed for persons beyond or older than the typical high school student.

Programs for preschool children through Grade 3 constitute 7 percent, while 12 percent are planned for children in Grades 4 through 6, or elementary school programs add up to 19 percent of the total of 936 programs being reported at this time.

2) *Subject matter breakdown*

Over 200 subject matter categories were employed in the Master table; this meaningful, but unwieldy number of categories was reduced to 20 for this abstract. The 20 categories are presented in Table II.

---

[1] Copies may be obtained by sending the author 1) a 3 x 5 card stating name, address, title, institution and 2) a legal size self-addressed envelope.

[2] The publication of this completed research is anticipated for the summer of 1967. Research will include separate author, title, publisher indexes, and approximately 300 cross references. Literature will be mailed to all members of ERANYS.

[3] Jonathan W. Varty, "Criteria for Evaluating Collegiate Programed Instruction," *Research and Practice,* Report of the Fourth Annual Convocation of ERANYS, (November 18-19, 1963), pp. 117-118.

[4] Jonathan W. Varty, "Programed Self-Instruction for Adults," *Instructional Improvement Through Research,* Report of the Sixth Annual Convocation of ERANYS, (October 19-20, 1964), pp. 179-180.

TABLE I

PROGRAMS OF SELF-INSTRUCTION FOR ALL AGES

By School Level or Adult Status Target Population

Target Population Categories	N	%	N	%
A. Graduate students and other adults	149	16		
B. Undergraduate College students	198	21		
Beyond or older than High School students Sub-total			347	37
C. Senior High school students	211	23		
D. Junior High school students	199	21		
Junior and Senior H.S. students Sub-Total			410	44
E. Students in Grades 4 through 6	113	12		
F. Students in Pre-school through Grade 3	66	7		
Pre-school through Grade 6 Sub-total			179	19
GRAND TOTAL			936	100

Note: *Table I* has been divided into six tables, 1A-1F in the report distributed at the ERANYS Convocation in Albany, November 14, 1966.

TABLE II

PROGRAMS OF SELF-INSTRUCTION FOR ALL AGES

By subject matter categories

Subject Matter content or Areas	Programs	
	N	%
Mathematics: Arithmetic, Modern, Statistics, and Traditional	231	25
Science: Biology, Chemistry, Geology, Physiology, and Physics	163	17
Language Arts: English: Grammar, Literature, Remedial, Writing		
Reading: Comprehension, Remedial, Vocabulary	129	14
Economics: Business, Banking, Management, Marketing, Taxes, Trade	80	7
Social Studies: Geography, Maps, Political Science, Sociology	51	5
Industrial Arts and Vocational Education	39	4
Business Education: Bookkeeping, Clerical, English, Secretarial	33	4
Psychology: Behavioral, Personnel, Public Relations, Salesmanship	33	4
Modern Language: French, German, Hebrew, Italian, Spanish, Russian	30	3
Nursing Science	23	2
Art, Design, and Music	19	2
Games: Bridge, Chess, and Mathematical	19	2
Data Processing and Computers	17	2
Program Materials	15	2
Medical and Dental	14	1
Speech: Audio discrimination, Listening, Parliamentary procedure	11	1
Philosophy and Religion: Logic, Symbolic logic, Applied logic	9	1
Special Education: Handicapped	8	1
Library Science	6	1
Study Skills	6	1
TOTAL	936	99

These data indicate that 231 programs (25 percent) are in mathematics, 17 percent in Science, 14 percent in Language Arts, 7 percent in Economics, 5 percent in Social Studies, 4 percent *each* in Industrial Arts, Business Education, and Psychology; 3 percent in Modern Languages; 2 percent *each* in Nursing Science, Art and Music, Games, Data Processing, and Programed Materials; and a rounded off 1 percent (from .0064 to .0149) *each* in Medical and Dental, Speech, Philosophy, Special Education, Library Science, and Study Skills.

3) *Current Trends*

The data in Table III are designed to point up trends from 1962 to date. The number of programs for selected years are recorded in columns two, four and six for 1962, 1963, and 1965; these numbers are cumulated and recorded in column eight. Each number is converted to its percentage and recorded in column three, five, and seven, while the cumulated percentage is recorded in column nine. It should be noted that the number of programs recorded in column 10 and their percentages recorded in column 11, *were collected independently* of all other date in Table III.

TABLE III
TRENDS IN PROGRAMS OF SELF-INSTRUCTION FOR ALL AGES by Categories

1	2	3	4	5	6	7	8	9	10	11
Subject-Matter Category	Sept. 1, 1962[1]		Sept. 1, 1963[1]		April, 1965[1]		Cumulative[1]		Nov. 1, 1966[2]	
	N	%	N	%	N	%	N	%	N	%
Business Education	4	3	19	5	11	4	34	5	33	3.5
Economics	5	4	3	1	6	2	14	2	80	8.5
English	16	13	51	14	56	19	123	16	140	15
Foreign Languages	10	8	21	6	9	3	40	5	30	3
Mathematics	52	43	123	35	49	17	224	29	231	25
Science	24	20	69	20	70	24	163	21	211	23
Programing	2	1.5	9	3	3	1	14	2	32	3
Social Studies	7	6	16	5	22	8	45	6	51	5
Miscellaneous	2	1.5	41	11	65	22	108	14	128	14
Totals	122	100	352	100	291	100	765	100	936	100

Note: The number of programs for each category (recorded in the even numbered columns 2, 4, 6, 8 and 10) is divided by the total number for its respective year; the percent for each category is recorded in the odd numbered columns 3, 5, 7, 9, and 11 for its respective year.

[1] Data and categories as reported by the Center for Programed Instruction, *Programed Instruction Materials 1964-65,* New York, Teachers College Press, Introduction Fig. 1.

[2] Jonathan W. Varty, *Programs of Self-Instruction For All Ages, 1966,* to be published 1967.

An arbitrary plus or minus 1 percent was used to locate trends from April 1, 1965 to November 1, 1966 and the following should be noted: little or no change in English, Programing, Social Studies, and miscellaneous; a decrease in the percentage of programs in Business Education, Foreign Languages, and Mathematics; an increase in the percentage of programs in Economics and Science.

It is recognized that a change in the definition of a category could produce other generalized trends. More definitive data on trends in programed learning will be included in this research at the time of publication.

REPORT OF AN EVALUATION OF THE TEACHER EDUCATION PROGRAM OF THE SCHOOL OF EDUCATION OF LONG ISLAND UNIVERSITY

Richard P. Harmon, Long Island University

Problem

The problem of this study was to determine the value professionally experienced teachers who had completed the entire teacher education program of seven courses at Long Island University placed upon the contribution of the program as a whole and of each course to certain selected teacher practices.

Population Studies

The population studied consisted of 16 L.I.U. graduates who completed the teacher education program during 1949-60 and who had taught a minimum of two school years. Of these graduates who completed the program, 57.3 percent had taught two school years. Forty-two teachers had 2-5 years, 49 teachers had 6-10 years, and 25 teachers had 11-14 years of experience. Fifty-seven teachers taught kindergarten through 6th grade, 34 taught 7th through 9th, and 25 taught 10th through 12th grade.

Instrument

Each teacher was asked to evaluate each course on a rating sheet listing 31 teacher practices. These had been chosen on the basis of previous faculty and student judgment of the appropriateness of the practices as reasonable outcomes of teacher education program courses. Teachers rated each course on a scale of one to five, with 1 representing negative influences on each practice used by the teacher, 2 representing no discernible influence, 3 little influence, 4 considerable influence, and 5 great influence.

Data

The data were analyzed to determine the influence of:

a. the total teacher education program and individual courses on combined teacher practices
b. individual courses on each practice
c. length of experience and grade level of teaching on teachers' evaluations

The statistical procedures included analysis of variance.

Findings

1. All selected teacher practices taken together:
   A. Total teacher education program—rating: 3.07 or little influence
   B. Courses
      1. Student Teaching—rating: 3.42 or little influence
      2. Methods—rating: 3.25 or little influence
      3. Child Development—rating: 3.19 or little influence
      4. Seminar—rating: 3.14 or little influence
      5. Educational Psychology—rating: 3.04 or little influence
      6. Problems in American Education—rating: 2.87 or no discernible influence
      7. History and Philosophy of Education—rating: 2.57 or no discernible influence
   C. Influence of Variables
      1. When the teachers' evaluations were grouped according to years of experience those with 2-5 and 6-10 years of experience gave the total program significantly higher ratings than did the teachers with 11-14 years of experience.
      2. When the teachers' evaluations were grouped according to the teachers' grade levels of instruction no statistically significant differences appeared.
2. Selected teacher practices considered separately:
   A. Only three courses of the teacher education program were rated as exerting considerable influence on individual teacher practices. Student teaching and Methods had considerable influence in the use of multiple texts, while Student Teaching and Child Development had this level of impact in understanding pupil behavior. Student Teaching alone had considerable influence on two practices: arrangement of material for sequential learning and insistence upon good work habits. Child Development alone had considerable influence upon acceptance of every pupil.
   B. Five courses of the teacher education program were rated of little influence on a majority of

the individual practices. Student Teaching and Seminar exerted little influence on 25 practices. Also, Methods, Child Development, and Educational Psychology showed little influence on 23, 20, and 16 practices, respectively. There were five practices among these on which all five courses were evaluated as having had little influence. These were: the stimulation of participation in co-curricular activities; the provision of leadership opportunities; the provision of activities designed to develop cultural appreciation of ethnic groups: the use of diagnostic procedures; and the evaluation of one's own teaching in the light of the school's philosophy of education. Problems in American Education was rated of little influence on ten practices and the History and Philosophy of Education of little influence on one practice.

C. The courses were evaluated as having had no discernible influence upon the number of practices indicated in the following list: History and Philosophy of Education, 30 practices; Problems in American Education, 20 practices; Educational Psychology, 15 practices; Child Development, 9 practices; Methods, seven practices; Seminar, six practices; and Student Teaching, two practices.

D. None of the courses of the program was evaluated as of great influence or of negative influence on any of the 31 teacher practices.

E. The effect of the experience and grade level variables on the teachers' evaluations of course influence on separate practices was found to be, with some exceptions, little or non-discernible. These exceptions were:

   1. Teachers with 2-5 years of teaching evaluated Educational Psychology as having considerable influence on three practices and Child Development and Student Teaching as having considerable influence on two practices.
   2. Teachers of kindergarten through sixth grade evaluated Educational Psychology as having considerable influence on two practices and Student Teaching as having considerable influence on one practice.

Conclusions

1. The teacher education program of Long Island University has had a "little influence" effect on the teaching practices of its alumni.

2. Laboratory and field experience courses were rated more influential on teacher practice than foundation courses.

3. The variation in the evaluations of the influence of the program and individual courses by the three experience groupings indicated that whatever this influence, it was not significantly altered during the 3rd through 10th year of teaching. There was a decline in program and course influence after the 10th year of teaching.

4. The limited variation in the evaluations of the influence of the program and the individual courses by the three grade-level groupings indicated that the impact of the program and courses was essentially the same on teacher practice at the elementary, junior high, and senior high school level.

5. The teacher education program is in need of careful examination to increase the influence on teacher practice of the History and Philosophy of Education courses and to extend the strengths of the Student Teaching, Seminar, and Methods courses.

A MONTE CARLO MODEL FOR THE PREDICTION OF PUBLIC SCHOOL ENROLLMENTS

Mary Griffin & John Schmitt, Boston College

The Monte Carlo Method, developed at the Los Alamos Research Center during World War II, uses random numbers to determine which of the many possible values of an uncertain variable is to be used in conducting a single simulation of a process, procedure, operation or technique. This report describes an application of the Method to the problem of school enrollment projections.

The effect of enrollment predictions on the future direction and quality of educational programs has been noted by many writers. The major difficulty with existing methods has been their failure to account for factors which can only be specified in probabilistic terms. It has also been claimed that the reliability of forecasting systems defies determination. Through simulation these limitations can be overcome, and in addition, the probable pupil-population effects of contemplated policy changes can be determined.

Parametric input to the system consists of the numbers of children by sex, in each preschool age group and each school grade, and the numbers of females in each of six child-producing age groups. For the remaining independent variables, estimates consisting of probable, high and low values are obtained, with the high and low estimates defining the 0.99 probability limits of the variable in question. These latter include proportion of eligible children to be enrolled in kindergarten; failure rates, by sex, for each grade; acceleration rates; dropout rates; incidences of emigration and immigration and associated family-composition characteristics; nonpublic school enrollment figures; birth rates within each female age group and changes in female population as a result of nonwife migration. For each year over which the forecast is to extend, a total of 148 triads (probable, high, low) of estimates is required, but rather than complicating the problem, these requirements illucidate the complications inherent in it and force a consideration of each.

In conducting a single iteration of the simulation for a particular year, each triad of estimates is solved as a beta distribution with a mode equal to the probable estimate and limits equal to the high and low values. The mean, then, is one-sixth of the sum of the limits and four times the mode, and the standard deviation is one-sixth of the difference between the limits. A random normal deviate is then multipled by the standard deviation and the product added to the mean to determine the iteration-specific value of the variable to be used in the simulation.

Proportions of migrating families having specific numbers of children are determined through similar procedures, and the iteration-specific number of children for *each* migrating family is then found by drawing a random number from a uniform distribution between 0.000 and 0.999 and assigning the number associated with that particular proportion. Random numbers are similarly employed to determine the sex and age or grade placement of each simulated child.

After determining iteration-specific numbers of births, dropouts, accelerations, failures, nonpublic-school enrollments and nonwife female migrations, appropriate accumulations are performed, and the results are recorded for use as pseudo-parametric input data for the simulation of the next succeeding year. One hundred iterations of the simulation of population changes are performed for each year of the forecast, and upon completion of the last year's final iteration, the record of simulation outputs (enrollments by sex and grade) is screened to produce enrollment probability distributions for each sex in each grade during each year of the forecast. Such output permits the school planner to judge, not merely that the probable male enrollment in grade three during 1975 will be 263, but that there is no more than a 0.05 probability that it will exceed 299, and that there is no more than a 0.20 probability that it will be less than 238 or greater than 289.

Use of the model without a computer would be unthinkable, and even with the fastest and largest machines, the present system will be quite expensive to operate. Simulation of migration effects is the most time-consuming aspect of the operation, and the writers are presently seeking a mathematical short-cut (through linear programming) as a means of removing this limitation.

The possible effects, however uncertain, of new developments, such as birth control pills, upon population forecasts are readily apparent. Simulation cannot reduce uncertainty, but it can expose and clarify the cummulative effects of interaction by a large number of uncertain factors over an extended period of time. Additionally, the requirements of this model make it practically impossible for the forecaster to ignore significant factors.

PERFORMANCE OF SCHOOL DISTRICTS ON FISCAL AND OTHER FACTORS IN RELATION TO AGENCIES OF BUDGET APPROVAL

Charles M. Bernardo, Central School Boards Committee for Educational Research

The relationships between school district quality and the arrangements for school board authority and economic responsibility are being seriously questioned throughout the country.

State constitutions have placed the responsibility for public education in the hands of the legislatures. The legislatures, in every state except Hawaii, have chosen to sustain a local agency rather than to operate the schools from a central board. School boards are thus state boards of local jurisdiction. As legislatures rarely, if ever, give their local arms free rein in the raising and spending of money, three divergent means have generally been employed to check the fiscal powers of local boards:

1. Public vote on the total budget or that part above a tax limit (let those who are the source of revenues approve expenditures).
2. Tax limit (the don't-go-beyond-this-point philosophy).
3. Fiscal dependence (let some more responsible and sagacious agency, a non-school governmental agency, approve the school board's budget).

The question is, under which budgetary approval procedures can school districts best provide quality education? Many different answers are coming from a variety of quarters. A recent report by the Regents Advisory Committee on Educational Leadership in New York State, for instance, includes among its numerous recommendations the following:

> The school board, like other legislative bodies of government, should be able to determine budget and tax rate without submitting them to an annual vote. The public should utilize with boards the same avenues of recourse exercised with other governmental bodies—the defeat and reelection of incumbents in succeeding elections. . . . It is felt that a limit of the taxing power of the school board is desirable; such safeguards as are now used in city school districts in New York State can be adopted.

This recommendation emerged from an investigation which employed a "reputational criterion." The committee presumably asked people of reputation what they thought about public vote and tax limitation in school districts and made their opinions the basis of the report.

Even the more outstanding writings on the subject of fiscal responsibility of school boards reveal a confused picture. Among nine studies reviewed intensively, three accord superiority to fiscal independence and two to fiscal dependence, while four conclude that the dependence-independence dimension makes no difference.

Investigations of the budgetary approval problem have not been truly comprehensive with respect to empirical analysis. A large scale study framed in an appropriate research model was needed. The Institute of Administrative Research, at the request of the Temporary Commission on City Finances of the City of New York has recently conducted such a study. More than 200 districts in the 50 states, currently being advanced funds from the Federal government, ranging from 3000 pupils upwards, were included in the investigation. Our problem was to find out whether education profited more when administered through an independent school board with its own budgetary approval powers or when, combined with other municipal functions, it is administered as one multifunction of government. The values of tax limitation and public vote also needed to be determined.

Performance was our criterion on fiscal and other factors. Public education is an economic institution in the sense that it is involved in fiscal competition for funds. Schools compete with public and private enterprise for personnel and material. Schools compete with the local government for tax resources. Indeed, school districts are even in competition with each other. So it seems logical to expect that the strengths and weaknesses of the various methods of budget approval would appear in financial data.

Fiscally independent districts showed up better than the fiscally dependent in almost every way.

The important measure of net current expenditure, the major cost factor, was higher in fiscally independent districts. The amount raised locally is a key part of the explanation. Dependency does, in fact, restrain the district so that it scores lower on this measure. Fiscally independent districts also gain a greater percentage of the total local revenue collected. Dependent school boards do less well because they must compete with government demands so much of the money raised goes to general government purposes.

Independent districts are also somewhat better off in the important measure of state aid. The general legislative intent behind fiscal dependence appears to be a corollary to the more niggardly attitude of state legislatures toward the dependent districts. The superiority of the independent districts is further upheld on measures of teachers' salaries, class size, and both specialist and clerical assistance.

Although independent districts without public vote tend to be superior to dependent districts, even they are considerably below the level of the independent districts with public vote. This observation applies to the fiscal position as well as to salary measures and the battery of quality related factors.

The evidence indicates that the laws under which legislatures require local boards to operate cannot be viewed for their effects in a unitary fashion. *One* characteristic or control, though important, is a component of a more meaningful pattern of regulations. Thus, the multivariate approach was employed to see how budgetary approval patterns and the dimension of school district size would behave in combination towards a measure of fiscal performance. The critical question asked at this point was, "Would districts within a variety of budget approval and size categories differ on the index of fiscal performance when wealth was taken into account?" Levels of statistical significance were obtained.

Effective Buying Income per Capita, the wealth indicator employed, as would be expected, predicts the Composite Fiscal Performance Index at a nearly perfect level of statistical significance (.00). The greater the wealth of a district, the greater is the likelihood of the district performing in a superior fiscal fashion. This relationship can be seen up to the income level of $6,000 to $6,500 per capita, after which the fiscal performance levels of the districts decline. The figures also show that the higher the income per capita of a school district, the slower the increase in fiscal performance.

After wealth has been statistically accounted for, the figures imply that the ideal situation is where the legislature gives its local agent, the school board, a relatively high degree of control over the source of funds (as through a provision of fiscal independence) and also tends to view the local electorate as a party to the decision (as through a provision for public vote). The opposite of this ideal is, of course, fiscal dependence. In either case where the legislature is reluctant to "go all the way," tax limits are enacted as restraints. A tax limitation proves to be a detriment to fiscal performance whether a district is fiscally dependent or independent. This observation is upheld at every demonstrated level of income per capita. The data show further that tax limitation is particularly deleterious in the absence of a public vote for exceeding it. The optimum mode of fiscal control insofar as the index employed is concerned, is fiscal independence without tax limitation. Fiscal dependence may be the better arrangement, however, under a handicapping tax limit.

The issue of school board fiscal control is acute in large school districts, especially in urban areas, but the effect of school district size on fiscal performance is not statistically significant either alone or in combination with budgetary approval patterns. We are currently doing further work on the question of size.

The matter of fiscal independence and dependence appears to be critical to fiscal performance only when the wealth dimension is statistically uncontrolled. After the effects of wealth are parcelled out, the apparent potential of the independent approach comes to fruition in the absence of tax limitation and the presence of public vote. Thus, the view that patterns of State control ought to be assessed in predicative combinations is borne out statistically.

Several recommendations emerge from the results of our investigation:

1. Judgements regarding the merits of fiscal independence and dependence should not be made without regard to the factors of school district wealth, tax limitation structure, and mechanisms for involving the public in the budget approval process.

2. Limiting the taxing powers of the school board should be avoided.

3. If a tax limitation structure is to be enacted, it should be flexible by providing for public vote in the district on an annual or regular basis.

4. In the absence of tax limitation or in the presence of a flexible tax limit, the fiscally independent procedure of budget approval should generally be looked upon with greater favor than the fiscally dependent procedure.

5. In the presence of an inflexible tax limit, (e.g., no public vote in the district, a constitutional amendment, action of the state legislature), fiscal dependence should generally be looked upon with greater favor than fiscal independence.

A STUDY OF TEACHER TENURE IN NEW YORK STATE AND SURVEY OF TEACHER TENURE LAWS IN THE UNITED STATES

William J. Hageny, State University College, New Paltz

THE PROBLEM

Part One

The purpose of Part One of this study was to survey the Teacher Tenure Laws of the United States to determine if any other states, with more recent statutory revisions, had evolved new and apparently better teacher tenure laws, regulations or procedures.

Part One of the study was therefore divided into two tasks. The first was to read all of the pertinent parts of the tenure statutes of all the states having Teacher Tenure Laws; the second to survey, by questionnaire, all of the executive secretaries of the various teacher associations in regard to their opinions about the strengths and weaknesses of the Tenure Laws of their states.

Part Two

Part Two of the study was an attempt to find out if there were common misunderstandings of the Teacher Tenure Laws of New York State. This was to be done by means of a questionnaire sent to a selected sampling of school board members, school administrators and teachers. It was a fundamental part of the working hypothesis of this study that there were deep and serious misunderstandings about the Tenure Law and these contributed greatly to much of the bitterness of present-day school personnel problems.

CONCLUSIONS AND RECOMMENDATIONS

Part One

1. The Teacher Tenure Law of New York State has not been revised in any substantial way in 21 years. This is so, despite the growing sophistication and militancy of teachers and teacher groups. It is recommended that the group of New York State statutes, commonly called the Tenure Law, be revised to remove present inconsistencies and ambiguities.
2. Prior to such a revision, it is recommended that the following ideas or procedures of other states on tenure be examined carefully:
   a. The inclusion of *all teachers* covered under the Tenure Law, now required in many states.
   b. The States of Washington and California have statutes that contain no probationary period—as we know it in New York State.
   c. The California Tenure Law that provides a panel of expert witnesses who may be called to testify on matters of professional competence.
   d. The Indiana statute provides that dismissal (except for immorality or insubordination) take place at the *end* of the school year.
   e. The statute of Michigan in that the teacher may be suspended pending the hearing *but his salary is continued.*
   f. The State of New Jersey that uses the local board as more of an "indictment" and leaves the judicial decision to the State Commissioner of Education, who holds a final hearing.
   g. In Minnesota—when the final decision is in favor of the teacher, the charges must be physically expunged from the record. (Also in Ohio and Pennsylvania).

Part Two

1. The surprising conclusion to this part of the study which was designed to test the hypothesis "that there is a considerable amount of misunderstanding about the Tenure Law," was the fact that there is a considerable amount of agreement in the findings, among board members, teachers and administrators on many of the basic concepts of tenure; therefore, the conclusion reached by the author of this study is that board members, teachers and administrators are not as far apart in their thinking on this subject as we would imagine. Once the stereotypes and the tired clichés about tenure are removed, it is felt that all three groups could cooperate profitably on a just revision of the tenure law.
2. The second surprising conclusion is that teachers, on the whole, are less informed on the facts of the Teacher Tenure Law than are board members surveyed.
3. The third conclusion is that there is a misunderstanding about the Teacher Tenure Law. There

are also some sharp disagreements as to how the law is being operated.

4. On the basis of these conclusions, it is recommended that:
   a. The New York State Tenure Law be studied jointly by a coordinating committee of teachers, board members and administrators in an effort to reflect the climate of the times. (If this is not done soon, it is a fairly easy assumption to make that the tenure law may soon be negated by collective bargaining agreements that will spell out more exactly terms of dismissal and appointment).
   b. That a fairly comprehensive and practical pamphlet be produced, again jointly by the three groups, to help all parties concerned do a better and fairer job on dismissal proceedings.
   c. That teacher preparation institutions be requested to include in the curriculum basic material on the Education Law of New York State to help beginning and experienced teachers understand the "ground rules" of their profession.
   d. That the National Education Association be requested to establish a committee or study on a national level leading to the proposal of a "model law" in teacher tenure.

NOTE

This study completed in the Spring of 1966, was sponsored by the New York State Council for Administrative Leadership (152 Washington Avenue, Albany, N.Y. 12210). A complete report of the study will be available from C.A.L.

AN EVALUATION OF TITLE III (NDEA) IN NEW YORK STATE

William T. Callahan, The Education Council for School Research and Development

Was Title III of the National Defense Education Act a boon or a bane to elementary and secondary education in New York State? To shed some light on this question, the State Education Department sponsored an evaluation[1] of the Act's impact on statewide and local efforts to improve instruction in science, mathematics and modern foreign languages. The major conclusions of this study are summarized in the following paragraphs:

1. Title III participation by local school districts did not significantly stimulate total expenditures per pupil, but did in some cases bring about a reorientation of expenditure patterns to the detriment of other subject matter areas.

2. Participation in Title III, measured on the basis of dollars per pupil, was quite uniform in the school districts of New York State. Examination of the relationships of two variables, school district wealth, and enrollment, to participation revealed only correlations of a low order. The much-quoted statement that spefic purpose, non-equalizing grants-in-aid favor the wealthier school districts (the rich get richer), was not borne out by this study.

3. Little opposition to Federal aid was revealed during the course of the evaluation of Title III. Only 95 school districts failed to participate in the Act during the period 1958-59 through 1963-64 and their nonparticipation was related to small school size, lack of financial and staff resources and other local factors.

4. Federal funds earmarked for the strengthening of State supervisory services seemed, on a dollar for dollar basis, to have a greater effect on the improvement of instruction than those funds made available to local school districts for the acquisition of hardware.

5. Analyses of enrollment trends, Regents examination papers written, and student participation and achievement in Advanced Placement programs as well as competitions such as the Westinghouse Science Talent Search, clearly demonstrated a heightening of student interest in science, mathematics and modern foreign languages during the Title III era.

6. Evaluation of the success or failure of their Title III programs by local districts was virtually nonexistent. In few instances were the objectives of the equipment acquisition programs clearly stated, and in no instance were objectives couched in behavioral terms.

While the evaluation of Title III yielded little hard data, many indications pointed to the fact that, in general, instruction in science, mathematics and modern foreign languages had been strengthened during the life of the Act, and it was logical to conclude that Title III played a part in the improvement process.

---

[1] Bureau of Department Programs Evaluation. National Defense Education Act—Title III—An Evaluation of the Program in New York State.

A DISCRIMINANT ANALYSIS OF VARIABLES USED TO SELECT STUDENTS FROM DISADVANTAGED BACKGROUNDS[1]

Roscoe C. Brown, Jr. & Edward H. Henderson, New York University

This study presents the results of an analysis of the variables used to select students for Project APEX, a college program for youth from disadvantaged backgrounds, supported by the Office of Economic Opportunity and the Astor Fund. Since the APEX program was designed for male students who were graduating from the general curriculum (a curriculum which excludes students seeking an academic, commercial or vocational diploma), it was necessary to identify two schools with an enrollment in the general curriculum that was large enough for a systematic process of selection to be used. It was decided to select the students from Morris High School in the Bronx and Benjamin Franklin High School in East Harlem in New York City.

The variables that were used in the selection process and the instruments used to provide data on them are listed below:

VARIABLES	*INSTRUMENT*
1. Intellectual Functioning	1. School and College Aptitude Test (SCAT) developed by the Educational Testing to identify verbal and quantitative aptitude.
	2. Stanford Achievement Test (Advanced Battery) Paragraph Meaning, Arithmetic Concepts, Social Studies Study Skills.
	3. Review of school record.
2. Personality and Social Adjustment	1. Gordon Personal Profile.
	2. Minnesota Counseling Inventory (modification of MMPI for high school counseling purposes).
3. Personal Goals and Motivation	1. Life Planning Questionnaire (developed by Martin Hamburger for the Career Patterns Study, Teachers College, Columbia University).
	2. Personal data sheet.
	3. One hour structured interview with a counseling psychologist.

Approximately 100 students from each school were tested. This analysis is based on the data for 148 students, 78 in the selected category and 70 in the not selected group.

After tests were administered and scored, three members of the professional staff of APEX rated each student on the appropriate criteria using a rating form. The responses to a yes-no question, "Should the student be in APEX?", was used to develop the list of students to be selected. Where three "yeses" were recorded, students were identified as selected and where two "yeses" were recorded the students were placed on an alternate list in a priority order determined by the staff.

The results of a discriminant function analysis of variance, a technique which identifies the relative contribution of each variable to the differentiation of the selected and not selected groups, showed that the major variable differentiating the two groups was the recommendation of the psychologist-counselor interviewer. When all variables were considered, four variables out of the 32 variables (school, ethnic background, athletic participation, and interviewer recommendations) account for 60 percent of the factors that differentiate between the selected and not selected groups. The school and ethnic variables, however, should be discounted because they relate to sampling factors. When school and ethnic background are omitted, interviewer recommendations account for 44 percent of the differentiation between the two groups and athletic participation accounts for 23 percent of the differentiation between the selected and not selected groups. This finding is not particularly surprising since selection of college students from this type of high school population is essentially subjective. No other studies have identified those variables which might improve the identification of students from disadvantaged backgrounds who have the potential to be successful in college.

---

[1] Supported by a grant from CAP—Office of Economic Opportunity, The Astor Fund and New York University.

The results of this study suggest that more research is necessary to develop different techniques for selecting students from disadvantaged backgrounds for special college programs. For the *present,* the data suggest that a carefully worked out counseling-type inventory used with a group of students, initially recommended by their teachers, is probably the most helpful procedure to select disadvantaged students for a special college program. This study does not, however, provide evidence on the extent to which the selection criteria predict the relative success of the students.

THE CONTRIBUTION OF SELECTED VARIABLES TO THE PREDICTION OF ACHIEVEMENT IN A SUMMER COMPENSATORY EDUCATION PROGRAM FOR ENTERING COLLEGE STUDENTS FROM DISADVANTAGED BACKGROUNDS[1]

Edward H. Henderson & Roscoe C. Brown, Jr., New York University

This report presents data on the prediction of achievement of 60 disadvantaged male youth in a summer compensatory education program. This program represents the first academic experience for the students in Project APEX, a college program for youth from disadvantaged backgrounds.

Sixty students who had been selected for Project APEX, began their first instructional experience at New York University in a specially scheduled eight-week summer program conducted through July and August, 1965.

The summer program had as its primary focus the development of basic communications skills. The selection data collected in prior testing, and previous knowledge of the characteristics of disadvantaged youth, indicated the need for concentrated attention on the improvement of those basic reading, writing, and speaking skills that constitute the general area of language arts. A major focus of the summer program was to obtain a more detailed estimate of the specific language arts and academic skills deficits of the APEX students, so that more precise projections for the September instructional camp program, and for the regular academic year program could be made. The summer program was organized as a noncredit remedial and enrichment program.

The students made a gain of .5 grade equivalent in arithmetic concepts and a gain of 1.5 in paragraph meaning (reading) over the two months of intensive reading and language arts instruction. It should be noted that no instruction in mathematics was offered during the summer. The over-all achievement performance of the students ranges from a mean language arts grade equivalent of 7.7 to an average social studies grade equivalent of 9.5. The data indicate that the students at the end of the summer program were performing at the upper eighth grade level in most of the academic areas measured by this test. It should be pointed out, however, that it might be expected that the students would not perform particularly well on standardized achievement tests in view of their backgrounds and previous academic performance. The multiple correlations to determine the relative contribution of each pre-summer variable to the prediction of achievement at the end of the summer program on the Stanford Achievement Test reveal that the pre-summer academic performance variables (achievement and intellectual ability) make the greatest contribution to the prediction of summer achievement. The data support what might be expected; previous academic performance is the best predictor of future academic performance. Although other variables are involved in predicting the academic performance of students from disadvantaged backgrounds, the academic factors make the largest contribution to the prediction of achievement.

Academic performance variables contribute over half of the variance on predicting achievement in arithmetic concepts, paragraph meaning, and social studies. The *r's* for these variables were between .88 and .99. In the prediction of arithmetic applications, language, and science, the contribution of the academic variables accounts for approximately one-third of the variance in the prediction. In the case of spelling ($r = .81$) there is no particular variable that makes the greatest contribution to the prediction.

---

[1] Supported by a grant from CAP—Office of Economic Opportunity, The Astor Fund and New York University.

THE RELATIONSHIP OF RIGIDITY AND AUTHORITARIANISM TO THE ACHIEVEMENT OF COLLEGE STUDENTS FROM DISADVANTAGED BACKGROUNDS[1]

Dorothy Lander, Harry Ploski & Roscoe C. Brown, Jr., New York University

Dogmatism and rigidity are characteristics of personality that have been widely studied. This paper presents the findings of a study to determine the relationship between these variables and the achievement of students in Project APEX, a college program for youth from disadvantaged backgrounds.

The F scale, developed 1943-1950 purports to measure social rigidity and right of center authoritarianism. The D scale, fully developed in 1959, measures social dogmatism. Correlations between the D and F scale have consistently been found to be significant; the correlation in the present study was found to be .65.

The means for the students in the brighter instructional group (Group I) are similar to those found for the middle class and college population, while those of the lowest achievers (Group III) approximated the norms established for the working class population. A t-test between the means for the groups indicated that the students in Group III scored significantly higher than Group I (t = 3.31) at the one percent level of confidence, and Group II (t = 2.19) at the five percent confidence level. It is therefore suggested that the students in the better achieving group of students were similar to the college oriented population in terms of dogmatism.

Correlations were computed between the students scores on the F and D scales and their scores in the SCAT verbal, SCAT quantitative, and total SCAT scores. No statistically significant correlations could be found. (The r's ranged from –.45 to +.39.) However, the correlation of the total SCAT scores and D scale score of Group III (r = .32) was significantly higher than Group I (r = .36) (t = 2.06), and approached significance for Group II (r = .29) (t = 1.82). The correlation of the total SCAT scores and the F scale score of Group III (r = .35) was significantly higher than the correlation of these variables for Group II (r = –.39) (t = 2.26). The difference between the correlation of these scores for Group III (r = .35) and Group I (r = –.23) approached significance (t = 1.74). When the F and D scale scores were correlated with the students' scores on the Stanford Paragraph Meaning tests, the pattern of the results was found to be similar to those of the SCAT correlations. No statistically significant correlations were found (r's ranged from –.30 to +.41). The correlation between reading scores and F scale scores of Group III (r = .41) was significantly higher than that found for Group I (r = –.30) (t = 2.17).

During the semester the instructors in the program were asked to evaluate the students on their apparent motivations, their performance within the Apex group, and their projected college performance. Indexes were then devised for these ratings and then correlated with the F scale and D scale scores. Again no significant correlations were obtained (r's ranged from –.45 to .36).

When the Minnesota Counseling Inventory sub scales were correlated with F scores, no significant correlations were found (r's ranged from –.20 to .48). No significant correlations were found for the family relationship and conformity subtests of the Minnesota with D scale scores. However, correlations of the students' scores on the D scale with their scores on the other Minnesota subtests proved to be significant. Statistically significant positive relationships were found to exist between high dogmatism scores and scores on the emotional stability, leadership, and social relationship subscales of the Minnesota counseling inventory. A negative correlation was found with adjustment to reality. Since dogmatism and acceptance of external authority have been shown to be positively related we might suggest that among these disadvantaged students, highly dogmatic persons accepted the standardized norms of behavior and exhibited a greater degree of emotional stability as a result.

The data support Rodeach's contention that highly dogmatic persons are unable to view reality with any logical inner consistency, and exhibit a need to reject aspects of the environment which threaten the core

---

[1] Supported by a grant from CAP–Office of Economic Opportunity, The Astor Fund and New York University.

belief system of the person. The statistically significant inverse correlation between D scale scores and adjustment to reality scores would support this thesis.

This study suggests that within the population of disadvantaged youth who have graduated from the general curriculum in high school, performance on standardized tests is adversely affected by rigidity of personality and dogmatism. The F and D scales therefore may well be effective screening devices among disadvantaged students in the general curriculum to differentiate between students who are more and less able to achieve in a college program.

TEACHING DISADVANTAGED NEGRO CHILDREN TO READ: A SECOND REPORT ON THE CRAFT PROJECT[1]

Albert J. Harris, The City University of New York

The CRAFT Project is studying the results of two main approaches and four methods in teaching reading to urban Negro children in disadvantaged neighborhoods. A preliminary report was presented to ERANYS in October, 1965. The project at present is following the original children through the third grade, and children in a replication study are now in second grade.

The first-year results showed a slight advantage of the Skills-Centered Approach over the Language-Experience Approach, with the Basal Reader Method slightly ahead of the other three methods. An analysis of instructional time showed that this is an important factor which obscures the relationship between methods and results. At the second-grade post-test measuring point the Skills-Centered Approach still seems to maintain a slight advantage; this is tentative, however, because covariance corrections have not yet been made.

---

[1] United States Office of Education, Projects 2677 and 3246.

SEX DIFFERENCES IN THE DEVELOPMENT OF CONCEPTS OF SPACE, MATTER AND ENERGY IN COLLEGE STUDENTS

Leona K. Adler, City University of New York, Hunter College

Introduction

Science knowledge is cumulative and it is built upon certain basic concepts. It is assumed that students entering college have sufficient understanding of these basic concepts so that they may serve as a foundation for the learning of further concepts that are introduced in science courses at the college level and so that they may aid in the interpretation of the natural environment.

Somewhere along the way, however, it appears that boys and girls part company in their understanding of basic science concepts, and by the time they reach college, the boys have outdistanced the girls in this regard. This finding was a byproduct in a study of the development of concepts of space, matter and energy in college students at the School of Education of New York University, completed in 1963.

The purpose of this investigation[1] was to determine the nature and the extent of changes in the understanding of concepts of space, matter and energy in college students, and to determine to what extent such changes are a result of instruction in three required science courses.

Analysis of the data gathered for this investigation suggests that the *course* of concept development as well as the *level* attained, may differ in men and women by the time of college entrance. Although these sex differences were found as a byproduct of the investigation not as the topic for central study, the investigator considers these accidental findings of interest as a fruitful area for further study.

Procedure

An objective test constructed by the investigator was administered to groups of students in the School of Education of New York University during the academic year 1962-63. This group included 269 students entering the School of Education in September, 1962, and 453 undergraduates registered in three general education science courses during the fall and spring terms of 1962-63.

The significance of differences between gains on the science concepts test before and after instruction in each science course, and before and after a period of maturation alone, was determined by means of the analysis of variance of gains.

Using the data from the entering freshmen, the investigator determined intercorrelations of scores on the science concepts test; the Verbal and Numerical parts of the Scholastic Aptitude Test of the College Entrance Examination Board; the Ohio State Psychological Test, Form 24; and the number of years of high school science instruction. Point biserial correlations were computed to determine the influence of specified high school science courses on the development of specific science concepts, and to determine the significance of sex differences in the scores of entering men and women on the science concepts test.

The Test

Test items were designed to test the understanding of taxonomic, static, dynamic and holistic aspects of a sample of subconcepts of space, matter and energy. Taxonomic concepts are those which relate to naming, or classifying or categorizing of instances into equivalence classes. Static refers to passive elements as distinguished from active elements of concepts. Dynamic concepts are those which pertain to change, or process, or operation. Holistic concepts are those which involve the comprehension of the relationship of constituent parts to a whole.

Despite the occurrence of borderline cases, it was possible to group the test items into four sections. Part I contains 48 true-false questions involving categorization. Part II contains 59 multiple-choice items that are for the most part descriptive or holistic. Part III contains 21 multiple-choice items that require a relatively great degree of "process" in order to select a correct answer. Another series of test items, 11 questions making up Part IV, was designed to measure verbalization of these concepts, the ability to select a correct verbal formulation of a concept, in contrast with items that test for actual understanding as shown by successful application of the concept.

---

[1] Leona K. Adler, *The Development of Concepts of Space, Matter and Energy in Students at the College Level* (Unpublished Doctoral Dissertation, School of Education, New York University, 1964).

The reliability of the major part of the test, Part II, which was used for group comparisons, was found to be .83 and the coefficient of correlation between scores of the same students obtained by repeated measurement was .87.

General Findings

1. The development of concepts of space, matter, and energy is influenced significantly by an orientation course in science and a course in physical science. There is no significant change in the understanding of these concepts during a semester without college science.

2. Scores on tests which measure the understanding of concepts of space, matter, and energy are significantly correlated with scores on tests which measure verbal aptitude and numerical aptitude. With years of high school science held constant, the coefficient of correlation between the scores of 245 entering freshmen on Part II of the science concepts test and their scores on the Verbal part of the Scholastic Aptitude Test was .52; for the Numerical part, the correlation was .45; and for the Ohio Test, the correlation was .28.

3. The development of concepts of space, matter, and energy is influenced by years of instruction in high school science as well as by specific high school science courses. Courses in high school physics contribute the most to the understanding of these concepts, and courses in high school chemistry make the next highest contribution.

Findings Relating to Sex Differences

1. At the time of college entrance, students show wide variation in their understanding of concepts of space, matter, and energy. Differences in scores on the science concepts test at the time of college entrance are associated with differences in (a) high school science background; (b) sex; (c) verbal and numerical aptitude.

2. Men students surpass women students in the understanding of science concepts, as indicated by scores on the science concepts test. The point biserial correlation for science concept test scores with sex, for the entering men and women students, was .423, which is significant at the 1% level of confidence.

Men students entering college have stronger backgrounds in physical science than women. As a group, men students were found to have completed more years of high school science, and a greater proportion of men had completed courses in physical science in high school.

The scores of entering men students surpassed those of entering women students, even when their high school science backgrounds and aptitude test scores were equated.

3. Verbal aptitude plays a more important part in the concept development of women than does numerical aptitude. With the exception of the test on dynamic concepts, all correlations between science concepts test scores and verbal aptitude test scores were higher than the correlations between science concepts test scores and numerical aptitude test scores. The situation was reversed in the case of men students. Numerical aptitude plays a greater part in the concept development of men than does verbal aptitude.

4. In the case of both men and women, numerical aptitude is more highly correlated with the development of dynamic concepts than is verbal aptitude.

5. The understanding of science concepts is more highly correlated with years of high school science instruction in the case of men than in the case of women.

6. The development of dynamic concepts is significantly correlated with years of high school science instruction in the case of men, whereas the correlation between these two factors is very close to zero in the case of women.

7. In the case of both men and women, descriptive concepts and verbalization of science concepts are influenced by high school science instruction more than are taxonomic and dynamic concepts.

8. A substantial number of students are concept resistant. Their concepts remain virtually unchanged by either high school or college science courses. Concept resistance is much more common among women than among men.

9. Men show consistent superiority in the development of certain types of concepts. Among these are: (a) concepts of relative distance; (b) concepts involving measurement, such as volume, density and pressure; (c) concepts of potential energy; (d) concepts of motion and of relative motion; (e) concepts of gravitation; (f) concepts of atomic structure.

10. Although there was no significant difference between the gains of men and women in scores on the science concepts test as a whole, before and after instruction in college science courses, the men did show consistent gains in the acquisition of certain concepts.

COGNITIVE AND INTELLECTUAL FACTORS IN A STUDY OF ENGINEERING STUDENTS

Henry L. Rosett, Herbert Robbins, & Walter S. Watson, Department of Psychiatry, The Mount Sinai Hospital, New York City & Cooper Union

The individual's capacity to work is of vital interest to many groups, including engineering educators and psychiatrists. Traditionally, a major criteria for judging the outcome of selection has been the academic average attained in engineering school by the students. Yet grades have little relationship to actual professional performance.

A shared interest in the multiple factors relevant to engineering careers led to a cooperative research study between the Cooper Union School of Engineering and the Division of Community Services of the Mount Sinai Hospital Department of Psychiatry. An eight-year longitudinal study of the entire 1962 graduating class was undertaken to examine the relationships between developing career patterns and psychological variables studied prior to graduation. These variables included cognitive and intellectual functioning, character structure, goals and values, and work style. This presentation will be limited to some of our initial findings on cognitive and intellectual functioning and academic achievement.

Recent psychological research has demonstrated individual consistencies on tasks not usually thought to be related, such as the ability to locate a geometric figure embedded in a complex design and the sophistication of body image when the subject draws a person. This behavior involves ego functions such as perceiving, recognizing, conceiving, judging and reasoning. The data has been interpreted as indicating the existence of broad organizational structures in the mind which account for these consistencies over a wide range of behavior. These organizational dimensions have been termed cognitive control principles. They offer a theoretical framework for investigating the important relationships between individual differences in styles of thinking, character structure, and work styles.

Three cognitive control principles were incorporated in this research: field articulation, scanning, and physiognomic perception. Field articulation represents a general ability to direct attention selectively without being distracted by competing opposing stimuli. Scanning measures the extensiveness of deployment of attention. Physiognomic perception is defined as a mode of perception in which objects are predominantly regarded through the motor and affective attitude of the subject.

Intellectual Variables: All of the subjects took the SAT Verbal and the CEEB Physics and Advanced Mathematics examinations as part of the admissions procedure. They were also given several tests which were not evaluated for admissions purposes. These included the Cooper Union Block Counting Spatial Relations tests and three sections of the Yale Educational Aptitude Test Battery, Verbal Reasoning, Quantitative Reasoning, and Mechanical Ingenuity. They were also tested on the Logical Analysis Device, a machine which assesses problem-solving ability. The criteria of academic achievement used were the high school average and four-year average at Cooper Union.

Results

A factor analysis was performed on the 19 cognitive and intellectual variables by means of a varimax rotation. The five factors derived were:

Academic achievement factor: This factor is defined by grades in engineering school and high school and on the CEEB Advanced Math. High factor subjects also gave more common associations to the Kent Rosanoff Word List, suggesting that high grades are partially related to the capacity to respond in conventional terms.

Abstract reasoning factor: Verbal tests, and those involving numerical and other forms of symbolic logic have their highest loadings on this factor. The extent of this ability cannot be inferred from engineering school grades.

Field Articulation factor: The high loadings of the Block Counting Spatial, Mechanical Ingenuity, and Logical Analysis Device on this factor, suggest that the capacity for selectively directing attention is important for complex technical work.

Scanning: High scanning subjects do better on the Yale Quantitative Reasoning test. Extensive deployment of attention seems to be an advantage when many bits of information must be integrated to derive general principles.

Physiognomic perception: This factor was not related to established engineering abilities, but may be manifested in future work styles.

Traditionally intelligence and ability are considered to be relatively isolated from broader personality functions. The theory of cognitive control facilitates integration of our knowledge of mental abilities with the psychoanalytic theory of personality. Cognitive control factors are more broadly defined than most ability factors derived from multivariate studies. They are not limited to consistencies on purely intellectual tasks, but also have a relationship to physiological responses and defenses against anxiety. Since they are not defined in terms of adaptive value, their utility depends on the particular task under consideration.

Cognitive control theory may help clarify what ability tests actually measure. When cognitive operations have been more clearly defined, their modification by teaching techniques can be studied. If individual differences remain fixed, they represent an important consideration for vocational guidance. Clarification of the relationships between cognitive style and character structure should also yield additional understanding of why some personality types excel in particular professional areas.

A METHOD FOR STUDYING VERBAL BEHAVIOR DURING PROBLEM SOLVING

Ralph O. Blackwood & Barbara Newman, University of Bridgeport

The "think aloud" method, introduced by Claparède (1917), Selz (Van de Geer & Jaspers, 1966), and Duncker (1935), presumably permits direct observation of thinking processes while Ss solve problems. This may be contrasted with the reports of thinking processes which result from introspection or verbal reports. Although the "think aloud" method is currently in frequent use, it does not produce consistently reliable data (Vinacke, 1952). The instructions, to do all thinking aloud, do not adequately control S's verbal behavior, and the resulting silent periods allow S to edit his thoughts.

An improved "think aloud" method, described here, used operant conditioning and automated equipment to insure continuous high rates of speech, thus allowing little or no time for editing or for covert mediating responses. However, it seems possible that "forced" thinking aloud at high rates may interfere with or distort normal problem solving processes. The purpose of the present study was to test the hypothesis that thinking aloud, under the improved method, increases errors and changes the rate of problem solution.

Method

Six volunteer undergraduate women served as Ss. The problems to be solved consisted of 150 two-digit multiplication problems. Electro-mechanical equipment was programmed to control verbalization rates by controlling the schedule of reinforcement. Points on a read-out digital counter were used as reinforcers; Ss were instructed to earn as many points as possible. S's microphone, voice-operated relays, and the behavioral research equipment, programmed the apparatus to detect any failures and to maintain high, stable rates of speech. Technically speaking, there were two concurrent schedules of reinforcement: (1) differential punishment for low rates (dpl); (2) differential punishment for long pauses (dpp). When the apparatus detected a failure in maintenance of the criterion speech rate (8 responses each 3.5 sec. and no pauses greater than .5 sec.), point losses at the rate of 2 to 5 points per second were indicated on S's "losses" counter. Ther experimenter rewarded *S* for correct solutions with 10 points on S's "reward" counter, or punished problem solving errors with 10 points on the "losses" counter, by pressing appropriate switches. The apparatus was also programmed to switch S from "thinking aloud" to silent thinking in alternate 5-problem groups.

In a preliminary training period, each S was coached for 15 or 20 minutes to avoid point losses by continuous talk, then given a booklet of 50 training problems. The purpose of the block of training problems was to eliminate any "warm up" or practice effects and thus to improve the probability of obtaining stable problem solving rates during the experiment. Following training, Ss attended 3 sessions, solving 50 problems per session, or a total of 150 problems. Half of these were solved aloud and half silently in alternate blocks of 5 problems; each S served as his own control.

Results and Discussion

The data did not support the hypothesis that the improved "think aloud" method increases errors or changes rates of problem solution. A difference score was computed for each S by subtracting the average number of seconds required to solve a problem silently from the average number of seconds required to solve a problem aloud. Time required to solve silently did not differ significantly from time required to solve problems aloud ($D_M = 0.2$; $S_{D_M} = 4.5$; $t = 0.08$). The data on frequency of errors in problem solving was slightly less clear-cut. Three Ss made equal numbers of errors in the aloud and the silent conditions, and 3 made more errors in the silent condition (specifically 4, 5, & 12 more errors). This suggests that thinking aloud may result in greater accuracy. However, a t test indicated that the difference in errors between the aloud and the silent method was not significant, ($D_M = 3.5$; $S_{D_M} = 4.3$; $t = 1.8$; $P > .10$).

In conclusion, the present study offers some evidence that the improved "think aloud" method does not distort normal thinking processes and consequently it implies that the silent periods which plagued the traditional "think aloud" method may be eliminated by means of operant conditioning and automated equipment. However, the evidence is weak since it is based upon the acceptance of a null hypothesis, a small number of subjects, and an experiment which did not control for interaction between the two conditions.

Bibliography

Claparède, E. "La psychologie de l'intelligence." Scientia. 22: 353-368. 1917.

Duncker, K. Zur Psychologie des Produktiven Denkens. Berlin. Springer. 1935.

Van de Geer, J. P. & Jaspers, J. M. F. "Cognitive Functions." *In* Annual Review of Psychology. Farnsworth, P., & others, *eds*. Annual Reviews Inc. 17: 145-176. 1966.

Vinacke, W. E. The Psychology of Thinking. New York. McGraw-Hill. 1952.

COGNITIVE DEVELOPMENT OF CHILDREN BETWEEN TWO AND SEVEN YEARS

Greta Morine & Ruth Formanek, Hofstra University

Jerome Bruner has stated that "much of our commerce with the environment involves dealing with classes of things rather than with unique events and objects. Indeed, the case can be made that all cognitive activity depends upon a prior placing of events in terms of their category membership."[1] If this thesis is accepted, then it would seem to follow that educators need to have more information about the categorizing ability of children.

One of the topics of particular interest to teachers in our courses, which deal with the cognitive development of children, has been the apparently rapid change in the capacity of the child to categorize. This is a capacity which is almost totally nonexistent in the young baby, but which often exists in adult-like forms at the age of seven years. How does the child learn to categorize? Is it something which is learned by the child on his own? How might children's discoveries be facilitated? It was questions of this nature which led to this study.

This was an exploratory study, carried out in the belief that it was important to observe what happened with different materials before norms were set up which might place a child in a position relative to his own age group. Therefore, the sampling was not random, and all age groups were not represented in the sample. The primary subjects were eight children, ranging in age from two years, no months to seven years, ten months.

The subjects were presented with a variety of materials, which included: a set of 16 squares of balsa wood, in different background colors, with abstract designs of simple lines and shapes; a set of 25 oak-tag squares, in different background colors, with a fish figure also presented in different colors; a set of 24 cards in three shapes, with boy or girl cut outs, on different colored backgrounds; a set of red and blue circles and squares made of oak-tag; a set of blocks, representing different items of furniture, as well as people, in different colors; a set of plastic farm animals in several sizes and colors.

Each of these materials could be categorized according to two or more bases. For example, the set of red and blue circles and squares could be grouped according to shape, color, or both. These particular materials were the same as those used by Inhelder and Piaget, but the manner of presentation and the instructions to subjects were different than those of Inhelder and Piaget.[2]

The directions for all materials used in this study were the same. In each case the subject was asked to "put together what goes together," "to sort," "to arrange," "to match," until the examiner believed the child understood what he was to do.

The reactions of all subjects were recorded on both film and tape. Although children's verbalizations were recorded, analysis of these suggested that performance alone was of primary importance. Most verbalizations were of an irrelevant nature. Therefore the study focussed upon the nonverbal observations recorded on film. One of the major advantages of the nonverbal procedure, of course, is that the films are impressive demonstrations for students studying child behavior and development.

The analysis of results was of a descriptive nature. Particular attention was paid to how children categorized, what apparent criteria they used, what types of materials appealed to them, and whether or not they finished tasks.

Using additional data gained from observations of children's responses to the materials in preliminary trials, tentative categories were set up. These were based on the assumption that realistic categorizing is a more highly developed type of relation to objects than playing. The ten categories of responses established were as follows:

1. no visible interest and/or refusal when offered objects; plays without interest
2. plays with interest; manipulates objects with or without a theme
3. some incidental categorizing; mostly playing
4. some categorizing with theme; some playing
5. mainly categorizing; little playing

---

[1] Jerome S. Bruner, Jacqueline J. Goodnow, and George A. Austin, *A Study of Thinking,* New York, Science Editions, Inc., 1962, p. 231.

[2] Barbel Inhelder and Jean Piaget, *The Growth of Logic in Young Children,* New York, Harper and Row, 1964, p. 203.

6. mainly categorizing with discontinuities or errors
7. categorizing in pairs only (with pairable materials), or in groups without apparently relevant criteria
8. categorizing according to relevant criteria but changing criteria frequently, often not finishing task
9. categorizing according to relevant criteria and carrying through until everything has been categorized; following a plan
10. interested only in creativity of designs; able to categorize according to relevant criteria

In considering these categories it is important to note that the performance of each child was related to the task at hand. A child might react quite differently to two different materials.

The difficulty of the task itself is an empirical matter, and gradations on the basis of difficulty can probably be made. (Difficulty here refers to the success or lack of success of the arrangement of materials for each task.) However, the preliminary evidence does not indicate that any one material is consistently easier for all subjects.

It is interesting to note that with the thematic materials, i.e., the furniture blocks and the plastic animals, responses were consistently toward the play level. This pattern seems to fit in with the findings of Bartlett and Bruner in their testing of adults.[3,4] Both found that concepts were formed more efficiently and accurately when materials were of an abstract nature, than when materials lent themselves to thematic interpretations.

These very preliminary indications suggest some interesting problems in terms of their implications for education.

The effect of the use of thematic or abstract materials upon children's ability to develop concepts in various subject matter fields may be a problem worth pursuing. Other problems for further study include the following: Can children's categorizing ability be improved by experiences in categorizing, beginning with simple, abstract materials and moving through a series of materials of graduated complexity? Would such experience help children to deal more abstractly with thematic materials? What is the normal range of categorizing behavior for children at various age levels?

The answers to these questions could probably provide educators with valuable information from which to derive implications for classroom practice. This particular study will be followed by a longitudinal study, using several of the same subjects. The plan is to observe changes in general categorizing behavior, as well as changes in subjects' reactions to particular materials.

---

[3] Sir Frederic Bartlett, *Thinking: An Experimental and Social Study, New York,* Basic Books, 1958, p. 30.

[4] *A Study of Thinking*, p. 103-112.

MODELS FOR CLASSIFICATION

Robert F. McMorris, State University of New York at Albany

Prediction of academic success has been the topic of thousands of studies. The models which will be discussed are most relevant to one aspect of prediction, namely classification, where classification may be termed the assignment of individuals to groups. Examples might be assignment of high school students to tracks, of college students to majors, of graduates to occupations. Distinctions will be drawn between two kinds of questions (and information), success and resemblance.

First, considering the estimation of success, one might ask questions such as "How successful would the student be as an educational administrator?", or "Would he be more successful as an administrator or as a researcher?" The typical statistical approach is that of regression. Here we would predict a success criterion (probably grades) from the best linear weightings of the predictors.

To add a step of sophistication, the probability of success may be considered. This step necessitates looking at an array of criterion scores for individuals with this set of predictors, asking what percentage of the array would be in the success region, and repeating this process for the various groups. The individual may then be classified according to the highest probability of success.

The probabilities of success may be computed as follows:

$P(S_i|M_i \cdot x) =$ probability of success in group i, given that the individual is a member of group i and has a vector of scores x.

$$= \frac{1}{\sqrt{2\pi}} \int_{z_i}^{\infty} e^{-t^2/2}\,dt$$

where $z_i = \frac{c_i - y_i'}{s_{y \cdot x}}$

where c_i = cutting point for success in group i

y_i' = predicted score on the dependent variable for group i

$s_{y \cdot x}$ = unbiased standard error of estimate for y based on the x vector

Turning to the question of resemblance, one might ask the extent to which an individual resembles members of an intact group. For example, is he similar to educational administrators? to researchers? To which group is he more similar? Various methods of profile analysis have been proposed to answer this question, e.g., Cronbach's D and Cattell's r_p.

A mathematically rigorous approach is that of probability of membership. For a given set of scores, the relative density is computed for each of the groups involved, as noted below.

$P(M_i|x) =$ probability of membership in group i given that the individual has a vector of scores x

$$= \frac{\frac{N_i}{|D_i|^{1/2}} \cdot e^{-\frac{1}{2}\chi_i^2}}{\Sigma \frac{N_j}{|D_j|^{1/2}} \cdot e^{-\frac{1}{2}\chi_j^2}}$$

where $\chi_i^2 = (Z - \bar{Z}_i)\, D_i^{-1} (Z - \bar{Z})'$

$Z = (Z_1, Z_2, \ldots, Z_r)$ –row vector of discriminant function scores

$\bar{Z}_i = (\bar{Z}_{1i}, \bar{Z}_{2i}, \ldots, \bar{Z}_{ri})$ –centroid of group i of the discriminant functions

D_i = variance-covariance matrix

$[y = \frac{N}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}X^2}$ –density of the univariate normal curve]

The third model considered (Tatsuoka, 1956) allows for the combination of the two types of information by multiplying together the corresponding probabilities of success and membership, specifically

$P(S_i \cdot M_i|x) =$ probability of success and membership in group i given that the individual has a vector of scores x

$= P(S_i|M_i \cdot x) \cdot P(M_i|x)$

To continue with our example, we could multiply the individual's probability of success in administration by his probability of membership in administration. We could find the similar success-membership product for research. Then we could classify him according to the higher product.

The fourth and final model considered here may be termed the sequential model. This model was originated by Dr. Donald L. Meyer. With this approach, the probabilities of membership are considered only if the probabilities of success are sufficiently similar to each other.

Referring to the formal statement which follows, if the probability of success for group i is greater than the probability for any other group by a difference of more than c, then the classification is made according to the highest probability of success, i.e., the individual is classified into group i. However, if the difference between the highest probability of success and any other is less than some arbitrary constant, then the probabilities of membership are used to determine into which of these groups the individual should be classified. He is classified according to the highest probability of membership when considering only the groups for which his probabilities of success were relatively high.

If $P(S_i|M_i \cdot x) - P(S_j|M_j \cdot x) > C$,
for all $j \neq i$, where c is some arbitrary constant, the individual is classified in group i.

If $C \geq P(S_i|M_i \cdot x) - P(S_j|M_j \cdot x) \geq 0$
for any $j \neq i$, for the subset of groups so defined, the individual is classified in group r, where $P(M_r|x) > P(M_s|x)$ for all $s \neq r$.

In conclusion, four models for classification have been examined. One model is based on predictions of success, a second on estimations of resemblance. Two other models provide for combining the two types of information. With any of the models a prescribed "institutional" strategy may be employed for classification, or the information could be integrated "clinically," for example, by the counselor and the student.

One intended outcome of this examination of models is that the questions we ask and the information we deal with may be viewed with more precision and with an additional context. Hopefully this increase in precision and context will be a help whether we are using a statistical or an "armchair" approach.

BIBLIOGRAPHY

Tatsuoka, M. M. Joint-probability of membership and success in a group: an index which combines the information from discriminant and regression analyses as applied to the guidance problem. Unpublished doctoral dissertation, Harvard Univ., 1956.

NONLINEAR PAYOFF FUNCTIONS

Kenneth I. Raymond, State University of New York at Buffalo

Cronbach and Gleser have presented a treatment of decision theory involving linear payoff functions.[1] This paper is an attempt to extend part of their theory to the nonlinear case.

The particular situation studied is an institutional decision, without a quota restraint, in which a person will be accepted or rejected on the basis of a single univariate measure.

A monotone increasing polynomial payoff function of degree M is assumed. Formulas are derived, using a predictor score y' as cutoff, showing total utility, gain in utility, and optimum strategy based on the number of men accepted.

The use of the formulas is demonstrated by means of artificial data.

---

[1] Lee J. Cronbach, and Golding C. Gleser, *Psychological Tests and Personnel Decisions,* Urbana, Ill., University of Illinois Press, 1965.

ON THE USE OF INTERACTION TERMS IN MULTIPLE REGRESSION EQUATIONS

Richard B. Darlington, Cornell University & Dieter H. Paulus, University of Connecticut

Researchers in the behavioral sciences frequently turn to multiple regression techniques as an alternative to the Analysis of Variance. Multiple regression techniques are, at times, preferred as they allow independent variables to be treated as genuinely continuous and, since statistical independence of independent variables is not required, all data on hand may be utilized. In multiple regression techniques interaction terms are expressed as simple products of linear terms. The standard form of a multiple regression equation which includes two independent variables and their interaction may be stated as:

$$\hat{Y} = b_1X_1 + b_2X_2 + b_3X_1X_2 + C.$$

The statistical significance of each of the b-weights can be evaluated by dividing the b-weights by their standard errors. The resulting statistics have a t-distribution. If b_1 and/or b_2 are found to differ significantly from zero, then there are said to be significant "main" effects; if b_3 differs significantly from zero, then there is said to be a significant interaction effect.

It is the aim of this paper to illustrate that if the weight received by the interaction term (b_3 above) does not equal zero, then the weights received by the linear terms (X_1 and X_2 above) are, in part, a function of the means of the linear variables, *unless* the linear terms have been adjusted to a mean of zero prior to the computation of the interaction terms. That is, contrary to all intuition, the b-weights of the linear terms in a regression equation, which includes interaction terms, are usually dependent upon the often arbitrary values chosen for the means of the variables. If, for example, two new tests are developed and their means are arbitrarily set at 50 and their standard deviation at 10 and they are entered as X_1 and X_2 in a regression equation of the type shown above, then generally *different* b-weights will be obtained than those which would be obtained if the means of both tests had been set at zero. This is true despite the well known and obvious fact that the selection of means has no effect on the correlations of the tests with each other or with a criterion variable "Y". Unfortunately, the standard errors of these weights do not vary proportionally to variations in the sizes of the corresponding weights. Different levels of significance may, therefore, be obtained for the weights of the linear terms almost at will. Further, the correlation between the interaction term and the dependent or criterion variable is also affected by the arbitrary choice of means of the linear terms. The implication of this in the use of stepwise multiple regression techniques is obvious.

The following set of equations tries to show why this happens. Suppose we start with a standard form of a multiple regression equation in which the scores on the independent variables "X_1" and "X_2", are expressed as deviation scores. Let $\bar{b}_1$, $\bar{b}_2$, and $\bar{b}_3$ be the weights in this equation. We have:

$$\hat{Y} = \bar{b}_1(X_1 - \bar{X}_1) + \bar{b}_2(X_2 - \bar{X}_2) + \bar{b}_3(X_1 - \bar{X}_1)(X_2 - \bar{X}_2) + C$$

Carrying out the indicated multiplications, we obtain:

$$\hat{Y} = \underline{\bar{b}_1X_1} - \bar{b}_1\bar{X}_1 + \bar{b}_2X_2 - \underline{\bar{b}_2\bar{X}_2} + \bar{b}_3X_1X_2 - \bar{b}_3X_1\bar{X}_2 - \bar{b}_3X_2\bar{X}_1 + \underline{\bar{b}_3\bar{X}_1\bar{X}_2} + C$$

Since the underlined terms are constants, they need concern us no further as they can be absorbed in "C." We have, therefore,

$$\hat{Y} = \underline{\bar{b}_1X_1} + \underline{\bar{b}_2X_2} + \bar{b}_3X_1X_2 - \underline{\bar{b}_3X_1\bar{X}_2} - \underline{\bar{b}_3X_2\bar{X}_1} + C'$$

Combining the first and third underlined terms and also the second and fourth underlined terms of this equation we obtain:

$$\hat{Y} = (\bar{b}_1 - \bar{b}_3\bar{X}_2)X_1 + (\bar{b}_2 - \bar{b}_3\bar{X}_1)X_2 + \bar{b}_3X_1X_2 + C'$$

The coefficients of "X_1", "X_2" and "X_1X_2" in this equation are the weights which would actually be calculated by a computer program used in the normal manner to find regression weights if "X_1" and "X_2" are not first adjusted to zero mean.

It can now be seen that if $b_3 \neq 0.0$, then the weights received by "X_1" and "X_2" in this equation are no longer simply a function of $\bar{b}_1$ and $\bar{b}_2$. They are affected by "$\bar{b}_3\bar{X}_2$" and "$b_3\bar{X}_1$" respectively. These effects are zero under only two conditions; when $\bar{b}_3 = 0$, or when $\bar{X}_1 = \bar{X}_2 = 0$. Since a researcher includes an interaction term in a regression equation precisely because he is unwilling to assume that the weight of the interaction term is zero, he must adjust the linear terms to zero mean (that is, set $\bar{X}_1 = \bar{X}_2 = O$) before the interaction term is calculated. Thus he can be assured that the weights received by the linear terms reflect a quantity analogous to a "main effect" in Analysis of Variance designs. It should be noted at this point that if three-way interactions are to be included in the equation, then all two-way interaction terms must be standardized as above. Otherwise the weights received by the two-way interaction terms will be distorted in the same way as the weights received by the linear terms were distorted above.

The reasons for the change in the correlation between the interaction term and the criterion variable may best be shown by means of an example using simple imaginary data. This example consists of the scores of four subjects on each of four variables. In order to simplify the example, let us assume that $\overline{X}_1 = \overline{X}_2$ which in turn equals some parameter "a". The left half of Table I presents the scores of the four subjects on a criterion variable "Y", two predictor variables "X_1" and "X_2", and the interaction term "X_1X_2". The right half of Table I presents the corresponding deviation scores.

TABLE I
SCORES OF FOUR SUBJECTS ON FOUR VARIABLES

	Absolute Scores				Deviation Scores			
Subject	**Y**	X_1	X_2	X_1X_2	**Y**	X_1	X_2	X_1X_2
A	-4	$-2+a$	$-2+a$	a^2-4a+4	-4	-2	-2	$-4a+5/2$
B	0	$1+a$	$-1+a$	a^2-1	0	1	-1	$-5/2$
C	0	$-1+a$	$1+a$	a^2-1	0	-1	1	$-5/2$
D	4	$2+a$	$2+a$	a^2+4a+4	4	2	2	$4a+5/2$
Sums	0	$4a$	$4a$	$4a^2+6$				
Means	0	a	a	$a^2+3/2$				

The first three columns of scores are given; the fourth column is the product of the previous two. The sums and means of these four variables are shown at the bottom. Subtracting the mean from each of the four scores in a column gives the entries in the right half of the table, labeled "Deviation Scores".

We can use these deviation scores to compute the covariance between any two variables in the ordinary way, by summing the cross-products and dividing by N (i.e. 4). Doing this for every pair of the four variables gives the 4×4 variance-covariance in Table II.

TABLE II
Variance—Covariance Matrix of the Four Variables

	Y	X_1	X_2	X_1X_2
Y	8	4	4	$8a$
X_1		$5/2$	$3/2$	$4a$
X_2			$5/2$	$4a$
X_1X_2				$8a^2+25/4$

We can now apply the formula $r_{xy} = \dfrac{\text{Cov}\,(x,y)}{s.d._x s.d._y}$ to obtain the corresponding correlation matrix. (In order to simplify the calculations, we will obtain a matrix of squared correlations.) Table III presents this matrix of squared correlations.

TABLE III
Matrix of Squared Correlations of the Four Variables

	Y	X_1	X_2	X_1X_2
Y	1.0	.8	.8	$\frac{a^2}{a^2 + 25/32}$
X_1		1.0	.36	$\frac{4/5a^2}{a^2 + 25/32}$
X_2			1.0	$\frac{4/5a^2}{a^2 + 25/32}$
X_1X_2				1.0

Referring back to the first three columns on the right side of Table I, notice that the criterion variable is an *exact* linear function of "X_1" and "X_2"; the "Y" column is simply the sum of the next two. Nevertheless, the validity of the interaction term (whose square is shown in the upper right hand entry in Table III) is greater than the validity of either "X_1" or "X_2" when a^2 becomes greater than 25/8. Even more surprising, inspection of the upper right hand entry shows that it approaches 1.0 as a^2 approaches infinity. (Recall that in this example $a = \overline{X}_1 = \overline{X}_2$.)

In summary, we can now see that if multiple regression techniques are used as an alternative to the Analysis of Variance, then if interaction terms are included, we must standardize all linear variables prior to the computation of the interaction term in order to avoid introducing a distortion in the size of the b-weights received by the linear variables. If higher order interactions are used, the lower order interactions must also be standardized to means of zero.

ITEM SAMPLING

Robert Barcikowski & Carmelo Terranova, State University of New York at Buffalo

In a recent article by Lord, the question is posed as to the least number of items and the least number of subjects necessary to provide adequate test norms. The concern here is only with the least number of items. It is assumed that obtaining subjects is not a problem —i.e., the concern is with testing time rather than with obtaining subjects. This appears to be a logical assumption in most educational and industrial settings —all the subjects are usually available but only a limited amount of time is allocated for testing.

Consider a population of people, a population of items, and a distribution of test scores dependent upon them. This paper describes a random sampling experiment, programmed on the IBM 7044 computer, concerned with obtaining the least number of items necessary for estimating the mean and variance of this distribution of test scores.

Item discrimination, item difficulty, and the range of ability of subjects were taken into account in obtaining item trace lines for these subjects (Torgerson). From the item trace lines, a person by item score matrix was formed (Lord). The following assumptions were made—a unidimensional continuum of the variable of interest; subjects are distributed normally along the continuum; the items are dichotomous; the items are independent of each other at any given point on the continuum; a rectangular distribution of item difficulty and item discrimination.

Specific numbers of items, $i = 10(5)50$, were sampled and estimates of the population mean and variance obtained (Plumlee), (Lord). Using a test of relative efficiency (Mood & Graybill), these estimates were compared to the estimates of mean and variance obtained from standard group sampling procedures (using all items). An item sample size (under the above limiting assumptions) which gave a parameter estimate having less variability about the parameter than the estimate obtained from standard group sampling procedures, was found in each case.

TABLE I

Discrimination Difficulty	Highly (.5-1.)		Random (0.-1.)		Poorly (0.-.5)	
	μ	σ	μ	σ	μ	σ
Rectangular	15	15	20	35	20	50+
Random (0.-1.)	30	50+	15	20	12	20

Table I reports the size of the item sample where the population parameters are more closely estimated by the items sample than by the group sample.

Consider that the item difficulty has been transformed to a unit normal deviate. The rectangular distribution of item difficulty, therefore, is restricted in range compared to the random distribution. In addition, the correlation between the item and ability level is greatest when highly discriminating items are used and least when poorly discriminating items are used (Richardson & Kuder). With this in mind, we can see that the highly discriminating condition is closest to "reality" (i.e., an item score directly related to the examinee's ability level). Note also that the poorly discriminating condition is really a chance condition and that the random condition of discrimination is a combination of the "real-life" situation and the chance situation. The expectations, then, are that the highly discriminating items will provide estimates of the parameters with the fewest items; more randomly discriminating items will be needed; and the greatest number of items needed to provide estimates of the parameters will be those that are poorly discriminating. This is borne out in Table I.

From Table I, it can be seen that in all cases, except for the rectangular distribution of items difficulty combined with highly discriminating items, the number of items needed to achieve a greater relative efficiency for the estimation of the population men is less than the number of items needed to achieve a greater relative efficiency for the estimation of the population variance. In the excepted case, the number of items needed is the same.

BIBLIOGRAPHY

Gulliksen, H.—*Theory of mental tests.* New York: John Wiley & Sons, Inc., 1950.

Keats, J. A.—*A statistical theory of objective test scores.* Melbourne: Australian Council for Educational Research, 1951.

______, & Lord, Frederic M.—"A theoretical distribution for mental test scores." *Psychometrika, 27,* p. 59-72.

Lindquist, E. F. *ed.—Educational measurement.* Washington, D.C.: American Council on Education, 1961.

Lord, Frederic M.—"An application of confidence intervals and of maximum likelihood to the estimation of an examinee's ability." *Psychometrika, 18,* p. 57-77.

______. "Use of true-score theory to predict moments of univariate and bivariate observed-score distributions." *Psychometrika, 25,* p. 325-342.

______. "Estimating norms by item-sampling". *Educ. & Psych. Meas., XXII,* p. 259-267.

Mood, A. M. & Graybill, F. A.—*Introduction to the theory of statistics.* New York: McGraw-Hill Book Co., Inc., 1963.

Plumlee, Lynnette B. "Estimating means and standard deviations from partial data: an empirical check on Lord's item sampling technique." *Educ. & Psych. Meas., XXIV,* p. 623-630.

Raiffa, H. & Schlaifer, R.—*Applied statistical decision theory.* Boston: Harvard Business School, 1961.

Richardson, M. W. & Kuder, G. F.—"The calculation of test reliability coefficients based on the method of rational equivalence." *J. of Educ. Psych., 30,* p. 681-687.

Torgerson, Warren S.—*Theory and methods of scaling.* New York: Wiley & Sons, Inc., 1965.

HEADSTART: EVALUATION WITH A NEW INSTRUMENT

Ruth Formanek, Hofstra University

The present study has two purposes: (1) to test a new instrument for its evaluative potential of preschool programs, and (2) to collect data which will permit its validation in the future.

Evaluations of preschool programs in the past have been made primarily intuitively or on the basis of intelligence and/or personality tests. The use of rating scales and interaction schedules for teacher behavior has been another source of data for evaluation.

The literature suggests that neither I.Q. nor personality tests have been found useful for evaluations. I.Q.'s are often shown to vary so considerably that pre- and post-testing becomes largely meaningless. Both rating scales and interaction schedules for use in preschools are in the experimental stage.

It was our aim to provide data for evaluation on a behavioral level and to demonstrate the differential effects of the program on children under different conditions or "settings." It is planned to follow these children for 2 years, at which time achievement tests should furnish answers to questions such as: Can school achievement be predicted on the basis of behavioral observations made in a preschool program? How does the Headstart sample differ from a non-Headstart sample after 2 years, i.e. can the Headstart summer program be shown to have had any effects on the children? What does the "behavior profile" look like for individuals, for groups, and how consistent over time is the behavior sampled? How consistently do the same behavioral traits in response to environmental changes manifest themselves?

The instrument described here uses specimen descriptions and time sampling as the methods of recording. All recorded observations are coded according to a system of categories based on the child's style of coping with his environment, a system first proposed by Lois Murphy and adapted to school settings by R. Spaulding with modifications by the author. (See Appendix.)

Each observation can be assigned its proper category number, thus permitting quantification and computer analysis. Categories deal with the child's reactions to his environment and permit observation of changes in behavior with antecedent changes in environment.

The categories covered include aggression, independence, sociability, dependence, conformity, passivity, and withdrawal, all referring to coping style. Reliability is established after a training period and is expressed in percent of observer agreement. In the present sample, reliability was over 90 percent.

The sample consisted of 58 Headstart children enrolled in an eight-week summer program operated by 4 private nursery schools in a town on Long Island. One nonparticipant observer was assigned to each of 5 classes in 3 of 4 nursery schools. Specimen descriptions were of 2 min. duration and were taken once a day for each child. Other data was collected or made available as follows:

Social worker's family history
Pre and post I.Q.'s: Stanford-Binet and Peabody
A new nonverbal sociogram based on time sampling
A sorting task
A "cognitive check list"

Results are, in this preliminary report, primarily normative. All settings are divided into those which are primarily "free play" and those which are "teacher directed" on the basis of the degree of authority and structure imposed by the teacher. All data is averaged separately for the two settings, separately for boys and girls, and separately for each of 3 time periods made up of 12 days each. Numerals refer to the average percent of time that observed behavior was recorded in each category.

"School adaptive" behavior was differentiated from "school maladaptive" behavior on the basis of percent of behavior in the respective categories and 2 extreme groups were formed for purposes of future prediction and correlation with school achievement. In addition, other extreme groups were also formed on the basis of I.Q., SES and aggressive or withdrawn behavior.

FOR PRESCHOOL AND KINDERGARTEN

APPENDIX

COPING BEHAVIOR I

C1 Grabbing of objects such as toys, food, personal belongings, etc.

Taking over such objects without permission.

Destroying the property of others.

Kicking, hitting, pulling, pushing, etc.

General Description: Aggressive Behavior

C2 Annoying and bothering other children by pulling at their clothes, tapping their shoulders or arms, touching their belongings, talking to them while they are doing something, etc.

Getting the attention of the teacher or other children by jumping up and down, calling out, making inappropriate noises, etc., solely for the pleasure obtained by being the center of attention (mainly during a TD setting).

General Description: This is somewhat less aggressive behavior than C1. This type of behavior would tend to get attention through negative means.

C3 Manipulating other children or the teacher so that they will do what the child wants them to do.

Directing activities and movements of other children.

Bossing or commanding them; taking over a play situation.

Enforcing rules during a teacher's absence: Telling other children what the rules are.

Temper tantrums: This type of behavior is, of course, not frequently encountered in a school at all.

setting and will therefore be coded rarely, if

General Description: Although this is "bossy" behavior, it may serve a useful purpose such as keeping order while the teacher is out of the room.

C4 Resisting doing what he is asked to do by teacher or peers.

Trying to maintain his own ways.

Does not cooperate with other children or the teacher.

Engages in his own activities regardless of what the class is doing—with the apparent intention of not having to conform to the desires or requests of others.

Defensive checking:

Child looks up at teacher after he has done something which he knows is wrong.

General Description: This is the behavior of a child who will not do as told because he wants to maintain his independence as far as possible. A key word here is "defensive." The self-directed activity is pursued as a means of defense, and not because the child is intrinsically interested in what he is doing.

C5 Working on a project or other productive activity by himself.

Playing with great interest on his own: building, drawing, painting, cutting, pouring, etc.

General Description: The child appears to be genuinely interested in his activity and concentrates on it. He is genuinely involved in a productive activity.

C6 The pupil reflects and ponders over something he has heard or seen. He listens intensely and pays rapt attention.

NOTE: Behavior which at first may be classified in category XI may later be categorized as VI if it persists.

General Description: This behavior can be described as "digesting" the information offered, either by the teacher or some other source.

C7 Sharing or contributing ideas and interests with other children or with teacher.

Helping with jobs in the classroom, e.g. volunteering information, wiping spilled paint, juice, etc.

Helping another child with work or play.

Offering ideas in a discussion. (This may be shown by calling attention to himself by standing or calling out, etc. This attention-getting behavior is to be distinguished from C2 which is solely to get attention. Here the child genuinely wants to contribute his ideas or interests.) The focus here is on contributions to ongoing classroom activities.

C8 This is similar to C7 but in addition there is a mutual give and take.

Working and playing together in a group or in pairs.

Tussling, friendly physical contact, "horsing around," would be categorized here.

Responding to social situations or games with pleasure and/or approval, such as laughter, clapping, jumping up and down, dancing, etc.

C9a Appropriate:
Needs help and tries to get it from teacher by asking or going to her.

C9b Wants attention from teacher
Shows her something
Runs after teacher

C9c Stays with teacher just to be close.
Imitates her movements
Wants physical contact
Clings to teacher, etc.

C9d Child does not want to separate from parent.
Cries when parent leaves.
Clings to parent, cries for parent after parent has left.

C10 Doing as he is told with no apparent enthusiasm or interest.
Submitting to the teacher's requests or requirements.
Conforming to routine procedures with no apparent interest.
Waiting for instructions or directions or biding time with no apparent interest in a self-directed activity.

General Description: This may be superficially similar to C5 but here (C10) there is no real interest in the work.

C11 Watches passively without great interest, at the same time keeps tabs on the activities of other children and teacher, looks at other children, looks out the window, etc.
Will be easily diverted from his activities by unexpected events, such as door opening, birds singing, movements of other children, etc.

General Description: Child looks at the speaker or some event but is not actively involved.

C12 Child pays practically no attention to either teacher or other children.
Daydreaming, gazing off into distance without looking at anything in particular, appears not to be fully awake, and he may be slumped in his seat or on the floor.
Squirming, wiggling, etc.

NOTE: Behavior is classified in this category when such behavior appears alone, e.g. if child is eating *and* rocking back and forth, such behavior is coded C10.

General Description: Child seems more in touch with own internal stimuli than with his environment. Such internal stimuli might include: itching, muscle cramps, tiredness, and his responses to them might include scratching, stretching, yawning, etc.

C13 Moving away from an unpleasant situation, hiding in a corner, not looking up when called upon, pretending he is not in the room, moving physically away from teacher and other children to where he is most inconspicuous, not answering when called upon, avoiding anything that might be unpleasant to him.

General Description: This behavior appears to be based on a desire to withdraw, i.e., not to be part of the situation. In the playground a child might be seen hanging around the edge of the playing field or the fringe of an activity—hoping to be left alone and not be required to participate.

CHANGES IN CHILDREN'S SCIENTIFIC CONCEPTS AS A FUNCTION OF OBJECT INTERACTION, SOCIAL MODELING, AND ATTENTION

John L. Reed, Skidmore College

This study reports the results of an experimental investigation of changes in children's scientific concepts as a function of object interaction, social modeling, and attention. Change in children's concepts was conceptualized as an aspect of cognitive restructuring requiring that the individual integrate percepts arising in object interaction or social modeling situations with his existing concepts. Cognitive restructuring was defined as an autonomous process instigated by cognitive conflict induced when overt or covert responses to a given stimulus situation are found inappropriate, thereby producing surprise or uncertainty. Object interaction was conceptualized as direct manipulation of objects related to the experimental task, and social modeling was conceptualized as changes in the observer's response patterns following observation of a model's performance. Attention was considered to be an aspect of stimulus selection which can be directed either by the *S* or by an external source.

A theoretical framework was developed by informally integrating aspects of Piaget's theory of cognitive development, Berlyne's theory of conceptual conflict, and Bandura's investigations of social modeling. A theoretical statement, based on the theoretical framework, was formulated and specific hypotheses were derived for experimental investigation. Specifically, the pivotal predictions were:

(1) Object interaction and social modeling situations in which attention is directed are more likely to induce cognitive restructuring.

(2) Social modeling situations induce greater cognitive restructuring than object interaction situations.

Subjects for the study were fourth-grade boys and girls from the school system of a small city in Northeastern New York State. One hundred subjects of approximately the same age were selected in order to control developmental level for the experimental task and minimize variability arising from perceived differences between the subjects and the models used in the experimental manipulations of social modeling. Subjects were selected at random and assigned to one of five experimental groups, each consisting of ten boys and ten girls to minimize possible differential effects due to sex. An analysis of variance performed on IQ showed no significant differences between the groups on that variable.

Each treatment condition consisted of a pretest, experimental manipulation of the variables, and a posttest and criterion task. Each *S* was interviewed individually in relation to the *S*'s conceptualization and explanation of the relationship between properties of an object and the amount of water the object would displace upon immersion. The pretest was designed to determine the *S*'s developmental level for the task, elicit predictions related to the amount of water each of four objects would displace, and determine the *S*'s explanation of the phenomenon.

Treatment conditions for manipulation of the dependent variables immediately followed the pretest and were designed to induce cognitive conflict and provide the *S* with opportunities for epistemic behaviors. Attention to the relevant variables, related to the water displaced by each object, was manipulated in one object interaction condition and all social modeling conditions by use of an overflow jar and rubber bands to note the amount of water displaced by each object. In one object interaction condition and one social modeling condition the *E* directed the *S*'s attention through questions related to the results of the *S*'s direct experiences with the objects or the model's performance. In a second social modeling condition, attention was manipulated by the model's concurrent explanation for his performance.

Conflict in the object interaction conditions was manipulated by allowing the *S* to perform experiments with the objects and a container of water. In the other object interaction condition, the *E* directed the order in which the *S* selected the objects for testing.

Social modeling was manipulated experimentally by allowing each *S* to observe the performance of a sixth-grade boy or girl who acted as a model. The models were trained to perform at the next highest developmental level for the task, and the model's attractiveness was controlled by maximizing factors selected on the basis of an empirical investigation of attributes making sixth-grade students attractive to students in the fourth grade. Half the boys and half the girls in each social modeling treatment saw a model of the opposite sex in order to minimize variability due to sex effects.

The modeling treatments began with the model's performance on the same pretest given each *S*. Responses of the model on the pretest were designed to induce conflict in the *S* and demonstrate correct responses to the interview questions. Following the pretest the model performed experiments to demonstrate that he had correctly ordered the objects on the pretest and that a correct principle had been used. The model used the overflow jar for the demonstrations, and the various amounts of water displaced were marked with rubber bands. A post-test to determine treatment effects was given to the *S*s immediately following each experimental treatment. Subjects were asked to order the objects again and to give the principle upon which the order was based. As a criterion task the *S*s were asked which of two objects, equal in size but different in weight, would displace the greater volume of water. Any response indicating a belief that equal-sized objects displace equal volumes of water was assumed to indicate cognitive restructuring.

Statistical treatment of the data indicated that the experimental manipulations had not produced significant differential effects in behavior on the criterion task. However, detailed examination of the protocol data revealed striking differences in the types of initial explanation given on the pretest that, had they been taken into account in the experimental design, would have afforded support for some of the predicted treatment effects. Further analysis of the effects of type of explanation on cognitive restructuring afforded support for the following prediction: social modeling situations induce greater cognitive restructuring than object interaction situations.

The significance of the present study lies in the fact that it opens a new avenue of research related to the effects of social modeling on cognitive restructuring. Such research may prove to have highly significant theoretical and practical implications.

THE INTERACTION OF SEX OF TEACHER AND SEX OF PUPIL AS A FACTOR AFFECTING TEACHERS' RANKINGS OF PUPIL ANXIETY

Donald J. Treffinger & Richard E. Ripple, Cornell University

A self-report anxiety measure was administered to 215 sixth-grade pupils. Pupils' scores were ranked and correlated with rankings made independently by their teachers (five men and five women). Correlations were obtained between test and teacher ranks for each teacher's entire class, and then separately for boys and girls. The correlations for same-sex rankings were more often greater than correlations for opposite-sex rankings. For each teacher, correlations tended to follow the pattern: same-sex ranking (highest), total class ranking (middle), opposite-sex ranking (lowest). Mean discrepancies between test and teacher ranks were significantly lower for male teachers than for female teachers, regardless of sex of pupil. There was also a significant interaction between sex of teacher and sex of pupil. Less mean discrepancy existed between test and teacher ranks when male teachers ranked male pupils, and when female teachers ranked female pupils, than when opposite-sex pupils were ranked. Results are interpreted on the basis of role theory, sex roles in American culture, and the nature of empathy responses.

PLANNING INSTRUCTIONAL MATERIALS FOR CERTAIN LOWER CLASS ELEMENTARY SCHOOL CHILDREN

George Movesian, Hunter College, City University of New York

Purpose of the Study

The purpose of this research project was to investigate the possibility that, given a free choice of materials, elementary school teachers would not differentiate between materials judged appropriate for lower class children and materials judged appropriate for middle class children.

Hypotheses

A) When given a free choice, teachers in schools of predominantly lower class children select materials that are not significantly different from those materials selected by teachers in schools of predominantly middle class children.

B) When given a free choice, teachers in schools of predominantly lower class children select materials judged more appropriate for the background and experience of middle class children than that of lower class children.

Definitions

1. Lower Class Children—those children whose parents, when gainfully employed, are classified as unskilled.

2. Middle Class Children—those children whose parents, when gainfully employed, are classified as professional, proprietary, and managerial, except farm.

3. Appropriate Materials—those instructional materials which are significant to the learner. The materials which, when presented or referred to learners, are comprehended because of the learner's own background of experiences.

4. Resource Unit—a carefully planned series of suggestions centered in some broad problem, topic, or area of experience and organized to serve as a source of ideas, materials, and procedures to help a teacher in preplanning a learning unit.

Methods and Procedures

Subjects—28 and 29 teachers of lower and middle class children, respectively.

Selection Procedure—teachers were selected from a random list of schools identified as either lower or middle class by the respective superintendents, using Alba Edwards' *Alphabetical Index of Occupations and Industries* as a guide.

Instrument—a preselected resource unit titled "Transportation" for elementary school children. This resource unit was prejudged by a jury of three judges on the basis of which children's books were appropriate for children of lower and middle class background.

At a personal meeting with the teachers, the investigator explained the study and distributed the instrument and self-addressed, stamped envelopes. At their convenience, teachers were asked to select from the resource unit for their respective classes a teaching unit of two to four weeks.

The Chi Square Test of Association was used to test subhypothesis A, and the t-Test of Independent Samples was used for the other tests of significance. The 95 percent confidence level was used for all tests of significance.

Conclusions

The failure to reject subhypotheses A and B seems to support the possibility that, given a free choice of materials, elementary school teachers of lower class children do not differentiate between materials judged appropriate to be used by lower class children and materials judged appropriate to be used by middle class children.

It seems that teachers of lower class children are not meeting the slum children's needs of individual differences on a social class basis in the preplanning of instructional materials. The findings suggest that teachers of lower class children should re-examine the meaning, especially in relation to the disadvantaged, of the American ideals of equality of opportunity and human dignity. Equal opportunity to have an education seems to be lacking for lower class children when "equal opportunity" is defined as teaching children from the child's background of experiences.

EVALUATIVE STUDY OF PREKINDERGARTEN PROGRAMS FOR EDUCATIONALLY DISADVANTAGED CHILDREN

Louis T. Di Lorenzo & Ruth Salter, New York State Education Department

Preschool education for disadvantaged children is currently receiving unprecedented attention. In spite of the widespread institution of prekindergarten programs with generous State and Federal support, little has been done to ascertain the value of the preschool experience in counteracting the deficiencies in language development and in the perceptual and affective areas which handicap children of low socioeconomic background as they enter the regular school program.

This report summarizes the first-year findings with respect to scores on the Stanford-Binet Intelligence Scale of one evaluative study of Prekindergarten Programs for Educationally Disadvantaged Children. The Stanford-Binet was designated a criterion measure because I.Q. has been experimentally demonstrated and traditionally accepted as the primary indicant of capacity to learn.

Only five other studies of prekindergarten programs have been undertaken using the Stanford-Binet. Two of these, Alpern and Phillips, reported no significant differences between experimental and control children while three, Deutsch, Weikart, and Gray, indicated significant gains for disadvantaged children in prekindergarten programs. The approach to prekindergarten education was different in each of these studies.

The study reported here is a longitudinal one in which three successive groups of children will be followed into second grade, first grade, and kindergarten with attention being given to reading readiness, school achievement, and school adjustment.

The immediate findings are based on analyses of the pre-post Stanford-Binet scores of the first (1965-66) group of children. This 1965-66 sample consists of 569 children in seven school districts with 298 experimentals who attended prekindergarten programs and 271 who did not. The following criteria were used in the sample selection: Father's occupational rating according to the Warner Scale (disadvantaged = categories 5, 6, or 7) or general family status, e.g., being on welfare; age—3½ to 4½; eligibility for kindergarten in the following September; use of English in communication; and being toilet-trained. The children were pretested in June and July of 1965; matched in terms of sex, race, I.Q., and socioeconomic status; and then assigned at random to experimental and control groups in each district.

Each district was free to determine its own prekindergarten program for its experimental children. There were three major variations in program: four districts worked with homogeneous groups of disadvantaged children; two districts used heterogeneous groups with an admixture of disadvantaged and nondisadvantaged children as a program treatment; and the seventh district employed the ERE machine resulting from O. K. Moore's work in teaching infants to read. In the ERE district, the pupils came to school one hour per day, devoting 15 to 20 minutes to the machine itself, if interested. In the six other districts, the programs were half-day sessions which followed very closely the traditional nursery program. Language development aspects of the prekindergarten program, which were stressed in teacher workshops, were formalized to different degrees in different districts.

The goals of all the programs were those agreed upon by the research team representing the school districts (Greenburgh, Hempstead, Long Beach, Mount Vernon, Schenectady, Spring Valley, and Yonkers) and the State Education Department: (1) Increased capacity to learn, (2) greater language development, (3) better self-concept as a learner, (4) increased motor development, and (5) more positive attitudes toward school.

The post-testing of both the experimental and control children took place in June and July 1966. The bases for comparison in the analysis of the change in capacity to learn are the mean I.Q.'s of the experimental and control children classified by socioeconomic status, sex, race, type of program, and district. The findings derived from this analysis follow.

Comparison by Socioeconomic Status

1. The changes in mean I.Q. of the disadvantaged and nondisadvantaged experimental children and the nondisadvantaged control children were all negative, but none was significant.
2. The disadvantaged control group suffered a significant loss in I.Q. indicating progressive re-

tardation.

3. The I.Q. loss of the disadvantaged control group was significantly greater than that of the disadvantaged experimental group; the difference between the nondisadvantaged experimentals and controls was not significant.

Comparison by Sex

1. Both male and female control groups as well as the female experimental group experienced a significant loss in I.Q.
2. Only the experimental male group did not undergo progressive retardation.
3. The I.Q. losses of the four groups took the following order: greatest—control boys, second—control girls, third—experimental girls, and least—experimental boys.
4. The loss in I.Q. of the control males was significantly greater than that of the experimental males.

Comparison by Race

1. The nonwhite experimental group and both the white and nonwhite control groups demonstrated significant losses in I.Q.
2. The I.Q. increase of the white experimental group was significantly different from the losses of both the white control and the nonwhite experimental groups.

Comparison by Sex and Race

1. Only one of the eight subgroups, the experimental white males, made a positive change in I.Q.
2. The I.Q. change of the experimental white male was significantly better than that of both the experimental nonwhite and the control white male.

Comparison by District and Type of Program

1. Of the 14 district subgroups, 12 decreased in I.Q.; 4 decreases were significant.
2. In the seven district comparisons, only one experimental group, with homogeneous classes, was significantly better than its respective control group.
3. For all homogeneous groups combined, the difference between the increased mean of the experimentals and the decreased mean of the controls was significant.

In summary, it may be said that the prekindergarten programs did not enhance the I.Q. of the participants but rather prevented the extent of loss in I.Q. shown by the control children. The impact of the preschool experience was not to compensate for earlier deprivations but rather to inhibit further cognitive retardation. The overall effect of the programs was not to close the gap between the I.Q. scores of the nondisadvantaged and the disadvantaged as they entered kindergarten but rather to prevent the difference which existed at age four from increasing by age five.

Considering the experimental group as a whole, the prekindergarten experience was more effective for boys than for girls; it was also more effective for whites than for nonwhites. As regards the type of program and its effect on I.Q. for males and females combined, only programs with homogeneous classes improved capacity to learn.

These findings are the result of the first year of a three-year project in the schools and should in no way be regarded as final evidence. Further experience with prekindergarten education may produce different results.

BIBLIOGRAPHY

Alpern, Gerald D. "The Failure of a Nursery School Enrichment Program for Culturally Disadvantaged Children." Paper presented at the annual meeting of the American Orthopsychiatric Association, San Francisco, California, 1966.

Phillips, Leonard W., *et al.* "REACH (Raising Educational Aspirations of the Culturally Handicapped)." Unpublished report on Cooperative Research Project #5-8072-2-12-1, State University of New York College at Plattsburgh, 1965.

Institute for Developmental Studies. *Annual Report 1965.* New York: New York Medical College, 1965.

Weikart, David P., *et al. Perry Preschool Project Progress Report.* Michigan: Ypsilanti Public Schools, 1964.

Gray, Susan W. & Rupert A. Klaus. "An Experimental Preschool Program for Culturally Deprived Children." Nashville, Tennessee: George Peabody College, 1965. (Mimeographed.)

WHAT DO ELEMENTARY STUDENTS THINK OF SCIENCE FAIRS

Raymond G. Kenyon, Mid-Hudson Regional Supplementary Educational Center

Statement of Problem

Because of the changing philosophical and educational attitudes which have seemingly evolved toward elementary school science fairs, it was felt desirable to discover, examine, and interpret trends from 1962 to 1965 in the *fourth, fifth and sixth grade* MHSSC Science Fair participant attitudes, desires, feelings and educational values toward science fairs.

Analysis of Data

The data presented indicated that the status of fourth, fifth and sixth grade elementary school students toward science fairs was the same in 1965 as it was in 1962 on the basis of both empirical and statistical analysis. Children still reported what could be considered positive attitudes towards science fairs. They still indicated their high interest to enter another science fair (93%-96%[1]). They did not feel that the science fair entry was a waste of time (92%-97%). More student learning in the areas of arithmetic and mathematics took place due to completion of science fair project (76%-84%). Their feeling continued that more learning about science and mathematics accrued due to ther pursuit of science fair project than would have been available to them in the classroom (81%-86%). A home library aided them in information rsearch for their entry (75%-83%). This group of students who arrived at a regional science fair through competitive selection in their classroom, school, and school system, felt that pursuit of a science project was more worthwhile than many other activi-

---

[1] First figure in parentheses indicates 1962 percentage, second figure indicates 1965 percentage.

WHAT DO ELEMENTARY SCHOOL STUDENTS THINK OF SCIENCE FAIRS?
(A Three-Year Comparative Study of 4th, 5th, and 6th Grade Attitudes)

TABLE I	1962			1965		
	Yes	No	1	Yes	No	2
1. Would you like to enter a Science Fair next year?	93%	7%	—	96%	3%	1%
2. Did members of your family help you build, or complete your project?	27%	73%	—	24%	75%	1%
3. Did your teacher help you build or complete your project?	10%	90%	—	2%	98%	—
4. Did a neighbor or scientist help you build, or complete your project?	3%	99.7%	—	3%	97%	—
5. Do you feel that planning, constructing, and completing a Science Fair project was a waste of your time?	8%	92%	—	3%	97%	—
6. Did you have to learn more about arithmetic and mathematics to complete your project?	24%	76%	—	16%	84%	—
7. Do you feel that by competing in a Science Fair you learned something new about the use of tools; such as hammers, saws, drills, scissors, screwdrivers, and pliers?	52%	48%	—	55%	45%	—
8. Did your teacher show you how to use tools to complete your projects?	5%	99.5%	—	1%	98%	1%
9. Did your parents or friends show you how to use tools to complete your project for the Science Fair?	30%	70%	—	27%	72%	1%
10. Do you feel that by competing in a Science Fair you learned some science and mathematics that would not have been presented in your classroom?	81%	9%	—	86%	13%	1%

[1] No response—1962
[2] No response—1965

TABLE II	1962			1965		
	Yes	No	[1]	Yes	No	[2]
11. Did you have to use books or magazines from your library to find information about your Science Fair project?	75%	25%	—	83%	16%	1%
12. Did you have to use a school or community library to find information about your Science Fair project?	52%	48%	—	55%	45%	—
13. Do you feel that preparing a Science Fair Project is a more worthwhile activity than preparing for a school play, for example?	93%	7%	—	93%	7%	—
14. Do you feel that the rules used to judge the Science Fair; creative ability, scientific thought, thoroughness, skill, clarity and dramatic value are practiced in your classroom when you are learning science and mathematics?	82%	17.99%	1%	77%	22%	1%
15. Do you feel that the rules used to judge the Science Fair; creative ability, scientific thought, thoroughness, skill, clarity, and dramatic value are practiced in your classroom when you are learning social studies and language arts?	62%	37.98%	2%	59%	38%	3%
16. Do you feel that it is a good idea to hold the area Science Fair at the College at New Paltz?	97%	3%	—	94%	5%	1%
17. Would you rather have the area Science Fair held at another school or in a state armory?	11%	88.97%	3%	9%	88%	3%
18. Do you feel that the Science Fair at the College in New Paltz was well organized for you?	97%	2.99%	1%	91%	6%	3%
19. Do you enjoy the competitive nature of a Science Fair?	99%	1%	—	95%	4%	1%
20. Did you feel that you needed help with your presentation to the judges?	15%	84.98%	2%	11%	88%	1%

[1] No response—1962

[2] No response—1965

ties related to education, such as school plays (93%-93%). And an unexpected overwhelming percent of students enjoyed the competitive nature of the science fair concept (99%-95%), while at the same time they enjoyed the motivation of attending a regional science fair being held at a college (97%-91%).

When elementary school students were asked questions about who motivated them to undertake and complete their science fair projects, they responded that members of their family did not help (73%-75%); moreover, their teacher did not help with the project (90%-98%). They indicated that neighbors and/or scientists did not help them construct their projects (99.7%-97%). Over half of the participating students (52%-55%) felt they had learned more about the manipulative skills of using hand tools, while at the same time admitting that (99.5%-98%) did not receive help in these skills from classroom teachers. The percentage diminished to (70%-72%) responding negatively to receiving help from parents and/or friends in the use of tools such as hammers, saws, drills, scissors, screwdrivers, and pliers.

In summary the statistical analysis indicates that fourth, fifth, and sixth grade students entering the Mid-Hudson School Study Council's Regional Science and Mathematics Fair are self-motivated for this competition. They did not directly seek adult help to complete their projects. The students also felt that their self-directed science activity was not a waste of their time. They learned more about science, mathematics and manipulative use of hand tools by developing a project for the Science Fair than they felt they could have learned in the classroom. It is interesting to note that these students felt that their home library was a better resource for researching their project than was their own school library.

Aside from the statistical and empirical evidence presented in this study much could be gained from a descriptive study of the interpersonal educational relationship between the student and the panel of three judges who interview him about his project. All of the judges hold degrees in some field of science from a college and university. It is at the point of the judges entering into free dialogue with the student, that the student really finds out whether he has, or has not, attained creativity, scientific thought, thoroughness, skill, clarity, and dramatic value in presenting his project.

EFFECT OF VARIED PRESENTATIONS OF LABORATORY EXERCISES WITHIN PROGRAMMED MATERIALS ON SPECIFIC INTELLECTUAL FACTORS OF SCIENCE PROBLEM SOLVING BEHAVIOR

Donald Nasca, State University College at Brockport

Earlier studies by this investigator and evidence from aptitudes research have led to the development of hypotheses contending that junior high students possess independent abilities that can be identified, defined and measured. It was further hypothesized that appropriate content oriented instruments, designed to measure such unifactor traits, would be more successful in detecting differences resulting from variables in educational experiences than commonly used achievement tests.

Two test batteries, one of 13 variables (table I) and one of 10 variables (table III) were factor analyzed to identify relationships of newly developed experimental tests with reference tests already demonstrating unifactor traits. Of the 4 content tests developed, 3 demonstrated characteristics of tests designed to measure independent abilities, (tables II and IV). Two of the tests are clearly within the 'Structure of Intellect' (SI) framework. The third factor-pure test did not appear to fit into the noncontent oriented structure developed by Guilford and associates.

The factor-pure tests were then used to evaluate 3 variations of an educational experience. A 700 frame linear program designed to promote understanding of physical science principles was administered to 245 eighth grade students in intact, randomly selected classes. One hundred students (Part.) performed 73 science experiments while progressing through the program. Seventy-five students (Read.) read descriptions of the experiments while progressing through the program and 70 students (Dem.) observed a teacher demonstrate the experiments during their progress through the program. Approximately 10 class periods were required to complete the entire program and accompanying activities.

An analysis of covariance was used to statistically evaluate scores from the 3 factor-pure tests with standardized reading test scores being used as the control variables. The (Part.) group scored significantly higher than either of the other groups on the noverbal sensitivity to problem test (table V). There were no significant differences in performance of the 3 groups on the vocabulary test or verbal sensitivity to problems measure (tables VI and VII).

It was concluded that active participation in experiments supporting scientific principles was superior for

TABLE I

MEANS, S.D., and Reliabilities of 13 Test Variables

Test Name	Means	Standard Deviation	Reliability
1. Apparatus	12.54	5.56	.63
2. Card Rotation	68.27	21.02	.40[1]
3. Cube Comparisons	9.81	2.96	.41[1]
4. Division	19.23	7.28	.59[1]
5. Hidden Figures	4.42	2.69	.14[1]
6. Hidden Patterns	33.60	10.43	.49[1]
7. Multi. and Div.	25.89	8.82	.52[1]
8. Non-Verb. Prob. Solv.[2]	6.87	3.11	.73
9. Sci. Vocab.[2]	12.49	2.92	.62
10. Sci. Spac. Orientation[2]	11.67	3.55	.84
11. Seeing Problems	12.82	11.31	.52
12. Seeing Sci. Deficiencies[2]	5.56	2.46	.73
13. Vocabulary	15.51	5.38	.35

[1] Experimental tests

[2] Obtained communality as lower bound reliability estimate

development of at least one phase of nonverbal problem solving ability. No such superiority of method was evident when verbal abilities were being measured.

In general the research project has further demonstrated the feasibility of developing unifactor test instruments for detecting precisely how limited variables in educational experiences affect student behavior.

TABLE II
Five Factor Structure Rotated Factor Matrix

	CMU (Verbal Comp)	CFS (Spatial Orientation)	EMI (Closure)	CFU (Sensitivity to Problems)	MSI (Numerical)	h^2
1.	**47**	01	−14	**44**	−05	44
2.	04	**56**	06	24	14	40
3.	18	**52**	21	15	22	41
4.	06	13	28	−02	**70**	59
5.	−04	17	11	30	11	14
6.	−01	24	−16	**63**	−00	49
7.	02	18	−03	−02	**70**	52
8.[1]	02	−04	20	**37**	−19	21
9.[1]	**52**	09	27	−09	03	35
10.[1]	**36**	−01	**45**	11	14	36
11.	−13	15	55	−28	22	47
12.[1]	30	17	**44**	−09	05	32
13.	**44**	15	**35**	08	12	35

Note: Decimal points omitted.
[1] Experimental tests

TABLE III
Means, S.D., and Reliabilities for 10 Test Variables

Test Name	Mean	Standard Deviation	Reliability
1. Hidden Patterns	50.07	15.60	.35[2]
2. Mechanical Reasoning	44.14	7.66	.44[2]
3. Nelson Si. Read. Comp.	49.63	10.14	.86
4. Nelson Si. Read. Vocab.	66.45	8.71	.85
5.[1] Non-Verb. Prob. Solv.	7.65	3.01	.65
6.[1] Sci. Spac. Orientation	9.02	3.25	.84
7.[1] Sci. Vocabulary	13.15	2.55	.62
8. Seeing Problems	18.72	6.57	.21[2]
9.[1] Seeing Sci. Defic.	5.55	2.30	.73
10. Unusual Details	5.97	2.25	.54[2]

[1] Experimental tests
[2] Obtained communality as lower bound reliability estimate

TABLE IV

Four Factor Structure Rotated Factor Matrix, 10 Variables

	Comp.	Sci.	Ach.	CMU (Verbal Comp.)	h^2
1.	17	26	**49**	11	35
2.	08	**64**	13	−08	44
3.	**58**	01	**50**	**44**	78
4.	**56**	07	**51**	**38**	72
5.[1]	03	**65**	16	10	45
6.[1]	09	**45**	−05	07	22
7.[1]	07	**32**	20	**49**	38
8.	17	05	15	**39**	21
9.[1]	**51**	**37**	07	12	42
10.	08	02	**69**	23	54

[1] Experimental tests

TABLE V

Nonverbal Problem Solving Test (H_1), Analysis of Covariance

		Control Variable				Dependent Variable		
		Vocabulary		Comprehension		Unadjusted		Adjusted
Group	N	M	S.D.	M	S.D.	M	S.D.	M
I (Part)	100	70.24	8.39	53.20	11.07	8.32	3.38	8.20
II (Read)	70	64.47	6.99	48.64	8.48	6.39	2.97	6.38
III (Dem)	75	62.12	7.21	46.31	9.06	6.35	3.07	6.51

Adj. F = 5.88
Significant at .01 level

TABLE VI

Seeing Science Deficiencies Test (H_2), Analysis of Covariance

		Control Variable				Dependent Variable		
		Vocabulary		Comprehension		Unadjusted		Adjusted
Group	N	M	S.D.	M	S.D.	M	S.D.	M
I (Part)	100	70.24	8.39	53.20	11.07	5.77	2.51	5.46
II (Read)	70	64.47	6.99	48.64	8.48	4.83	2.55	4.95
III (Dem)	75	62.12	7.21	46.31	9.06	5.39	2.40	5.69

Adj. F = 1.80 N.S.

TABLE VII

Science Vocabulary (H_3), Analysis of Covariance

		Control Variable		Dependent Variable		
		Vocabulary		Unadjusted		Adjusted
Group	N	M	S.D.	M	S.D.	M
I (Part)	100	70.24	8.39	13.87	2.64	13.34
II (Read)	70	64.47	6.99	12.54	2.39	12.75
III (Dem)	75	62.12	7.21	12.84	2.95	13.34

Adj. F = 1.40 N.S.

PREDICTING SUCCESS IN MODERN FOREIGN LANGUAGES

Harold F. Bligh, Suzanne Farren, & Geraldine Von Maluski, Harcourt, Brace & World, Inc.

INTRODUCTION

In 1958 Paul Pimsleur and his associates undertook a thorough search of the literature on language prognosis published during the preceding 30 years[1] to ascertain what variables, if any, had been isolated and defined. Two factors—intelligence and motivation—appeared to contribute to learning to read and write a foreign language.

Starting with the assumption that the audiolingual approach to language learning calls for specialized abilities other than those required for learning in the traditional sense, Pimsleur initiated a series of studies in an attempt to identify and describe the factors which might be related to success in language courses stressing all four communication skills.

On the basis of extensive investigations conducted at high school and college levels, he concluded that aptitude for learning modern foreign languages could be defined in terms of 3 main factors—verbal intelligence, motivation or interest in learning the language, and auditory ability[2] *The Pimsleur Language Aptitude* Battery was structured accordingly into 6 parts which assess differing aspects of these four factors:

Part 1:	Grade Point Average in academic areas other than foreign languages. The student reports the grades he last received in four school subjects.	16 points
Part 2:	The student indicates on a five-point scale his interest in learning a foreign language.	8 points
Part 3:	A four-option vocabulary test of word knowledge in English.	24 points
Part 4:	Language Analysis—ability to reason logically in terms of a foreign language.	15 points
Part 5:	Sound Discrimination—ability to learn phonetic distinctions and to recognize them in different contexts.	30 points
Part 6:	Sound-Symbol Association—an association of sounds with their written symbols.	24 points
	Total	117 points

The research project described here is concerned with the efficiency with which the Aptitude Battery functions for particular groups and for specific purposes.

DESCRIPTION OF THE RESEARCH PROJECT

Three studies were designed to yield information on the (1) reliability of the parts and the total Battery, (2) difficulty of each part for successive grade groups, (3) intercorrelations of the parts and the total, (4) correlations of the parts and total with various criteria of success, and (5) appropriate weightings of the parts of the Battery to achieve maximal accuracy in predicting the criteria. Studies 1 and 2 accommodated separate analyses for boys and girls.

Research Study 1—Concurrent Validity Study. In the spring of 1965, the Aptitude Battery was administered concurrently with the experimental editions of the Pimsleur Listening and Reading Comprehension Tests to approximately 3,000 students completing first-level courses in French and Spanish. Forty-one schools in 13 states participated in the study. Final course grades, determined independently of the test scores, were reported by 7 of the cooperating schools.

Research Study 2—Predictive Validity Study. At the beginning of the 1965-66 school year, the Aptitude Battery was administered to students beginning grades 7, 8, and 9. A total of 1,201 students from 5 schools in 5 states was included in the grade 7 sample; 979 students from 6 schools in 4 states in the grade 8 sample; and 1,765 students from 9 schools in 5 states in the grade 9 sample.

Lists of those students beginning French, Spanish, and German at each grade level, served as the basis for the followup study at the end of the school year.

---

[1] P. Pimsleur, L. Mosberg, & A. V. Morrison. "Student Factors in Foreign Language Learning: A Review of the Literature." *Modern Language Journal,* April 1962, 46, 160-70.

[2] P. Pimsleur, D. M. Sundland, & R. D. McIntyre. *Under-Achievement in Foreign Language Learning: Final Report.* The Ohio State University Research Foundation (Columbus), May 1963. (Abridged version available from Materials Center, Modern Language Association, 4 Washington Place, New York, N.Y. 10003).

Teachers' grades and their ratings of student proficiency in each of the four communication skills served as criteria. In addition, the Pimsleur Listening, Reading, and Writing Tests were administered to ninth-grade students completing a first-level language course. All criterion measures were collected independently of the Aptitude test scores.

Research Study 3—Reliability Study. Split-half reliability coefficients for Parts 3 through 6 of the Aptitude Battery were derived from the scores of the groups tested in Studies 1 and 2. A test-retest reliability study was conducted during May and June in two school systems, one each in Massachusetts and New Jersey. A total of 238 sixth-graders and 235 eighth-graders completed both administrations of the Battery.

DISCUSSION OF PRELIMINARY FINDINGS

Although data processing and analysis were still going on at the time this paper was presented, certain phases of the study had been completed. The summary of the findings to date are discussed below. Copies of the statistical table are available upon request to the authors.

1. *Reliability data* consisting of split-half coefficients for Parts 3 through 6 of the Aptitude Battery were reported for the groups tested in Studies 1 and 2. These coefficients ranged from .57 to .82 with a median of .76. All but two (.57 and .60) fell above .60, the limit cited by Guilford as sufficient for relatively short tests, included a battery from which a composite score is to be derived.[3] The split-half correlations were of about equal magnitude for boys and girls.

2. Since the Grade Point Average and Interest measures do not lend themselves to analysis by the split-half technique, the *reliabilities* of these as well as the other parts of the Battery were assessed in Study 3 by the test-retest approach. Coefficients of .96 and .98 indicate that student-reported grades are highly stable over a 2-week period.

The importance of school grades as predictors of success in subsequent school learning has been consistently substantiated by research. The necessity of checking into school records which may not be readily accessible, the lack of reliability of school grades, and difficulties faced in translating grades from different sources to a common base, have discouraged their use in the past. Predictors of this kind, however, can be conveniently built into an aptitude battery. Applications of such a technique are found in the *American College Testing Program*,[4] as well as in the Pimsleur Language Aptitude Battery.

Test-retest coefficients of .80 and .90 were found for Part 2, Interest; of .72 and .74 for the Verbal score (Parts 3 & 4); of .72 and .76 for Auditory score (Parts 5 & 6); and of .86 and .89 for the total Battery score. These values as well as the split-half coefficients seem commensurate with the test lengths and adequate for use of the Battery in selection and guidance. In assessing the usefulness of the Battery for diagnostic purposes, it should be noted that Parts 3 and 4 are combined to yield the Verbal score; and Parts 5 and 6, the Auditory score. Use of the scores from Parts 3, 4, 5 and 6 separately would not be recommended in light of the reliabilities reported here.

3. *Means and standard deviations* of part scores and total scores were reported separately by sex and for sexes combined for all groups tested, while standard deviations, and standard errors of measurement for the verbal, auditory, and total scores were noted for the test-retest groups. Although the data presented were cross-sectional in the sense that all of the participating schools had not tested at consecutive grade levels, it appears that the total Battery becomes easier with successive grade levels. However, growth in the auditory skills seems less pronounced than growth in the verbal skills. Further study will be directed to the question, "Do sound discrimination and sound-symbol association level off earlier in the learning process than vocabulary and word analysis or is this finding a consequence of the test content and structure?"

Relatively low mean scores on Parts 3 and 4 at the end of grade 6 and beginning of grade 7 suggested that those tests may be slightly speeded for these levels. Subsequent examination of the answer sheets supported this contention. The reliability coefficient for Parts 3 and 4 tended to be lower for these groups than for groups where Parts 3 and 4 were less difficult.

It was concluded that the Aptitude Battery is appropriate for end of year testing at grades 6 through 9. In addition there appears to be sufficient ceiling for each part to warrant its use in grades 10 through 12. The Auditory tests, Parts 5 and 6, would probably function adequately at end of grade 5. However, the Verbal tests, Parts 3 and 4, appear to be too difficult for most fifth-grade classes.

3

---

[3] J. P. Guilford. *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill Book Company, 1965.

[4] Leo Munday. Comparative Predictive Validities of the American College Tests and Two Other Scholastic Aptitude Tests. *ACT Research Reports.* Iowa City: Research and Development Division, American College Testing Program, August 1965, 6, 1-14.

4. *Data on the heterogeneity* of the Aptitude Battery were analyzed for various groups. Low to moderately low inter-coefficients, ranging from .10 to .56 for separate parts, and from .48 to .52 for correlations of the Verbal and the Auditory scores indicate that the abilities measured by the separate parts of the Battery are relatively distinct.

5. *Correlations of part and total scores* with several different types of criteria were derived for groups tested concurrently with the predictor and criterion measures and for other groups with prediction spanning a school year. The factors contributing to prediction differed from criterion to criterion and from situation to situation. To achieve maximal accuracy in prediction, appropriate weighting of the parts of the Battery should be determined for each group and criterion.

The total score as weighted by two scoring procedures reported in the test *Manual* is the single best predictor of the Pimsleur Proficiency Test, but is slightly less valid than student-reported grades in predicting teachers' grades. The efficiency with which students' reported grades predict teachers' grades was noted by Pimsleur in several of his experimental studies. He felt that this occurrence might well be due to the fact that foreign languages are probably taught much like other school subjects.[5] From the data reported here, it appears that the more audiolingual teaching and grading become, the less weight student-reported grades carry in predicting success. While student-reported grades correlated .56 with teachers' grades over a period of a school year, they correlated only .25, .18, and .20 respectively with standardized tests in listening, reading, and writing. Further analysis of the data will permit a more definite conclusion on the functions of self-reported grades and interest in predicting success in academic subjects.

Studies 1 and 2 provided data for empirically establishing the validity of the Pimsleur Language Aptitude Battery for predicting success in French, Spanish, and German at three grade levels and for a variety of criteria. These data are currently being processed. As a part of the analysis, coefficients of multiple correlations, and regression equations will be developed. Attention will be given to differential prediction in the three languages and for boys and girls.

---

[5] Pimsleur, et al. (1963).

SUBJECTIVE ANALYSIS OF AN EXPERIMENT COMPARING A TEAM AND A CONVENTIONAL APPROACH TO THE HIGH SCHOOL BIOLOGY COURSE

Paul Joslin, University of Rochester; John Montean, University of Rochester; John Schmitt, Boston College
David Robinson, Greece Central Schools

A subjective report is made at the end of the first year of a two-year controlled evaluation of a team approach to the high school biology course, involving about 1000 students and 17 teachers in 6 schools in the Rochester area. Some of the hypothesized advantages and disadvantages of team teaching are observed; some are not. Other observations are also made.

An objective report on the experiment will become available during the 1967-68 academic year. Other purposes of the study are to, (1) develop a basic model by which instructional innovations may be evaluated,[1] and (2) resolve barriers to the development of sound achievement tests by classroom teachers.[2]

Hypothesized Advantages Observed

1. Team planning provides practical and effective inservice education.
2. New teachers are successfully inducted into schools.
3. There is a greater potential for the best use of teacher talent.
4. Improvements may be made in correlation of subject matter topics and there is opportunity for the development of sequences of content and intellectual processes.
5. Group pressure stimulates better preparation on the part of teachers.
6. Audiovisual aids are used more creatively and efficiently.
7. Team structure facilitates flexible grouping.

Hypothesized Disadvantages Observed

1. Independent teachers are inhibited in teams and don't like working in teams.
2. Older teachers, both in age and experience, adapt to teams with difficulty. (New and younger teachers adapt with ease.)

Hypothesized Advantages Not Observed

1. Greater student interest.
2. Greater individual effort on the part of students.
3. Teachers are freed from routine tasks.

Hypothesized Disadvantages Not Observed

1. It is difficult to recruit teachers who can work in teams.
2. It is difficult to find strong team leaders.
3. Scheduling team classes is done only with great difficulty.
4. There are irritating effects on the existing departmental and grade organizations.
5. The morale of nonteam teachers is lowered.

Other Observations

1. Administrative assent and cooperation is absolutely essential.
2. Although not essential, specialized facilities and equipment are highly desirable.
3. Provision must be made for flexible scheduling.
4. Discipline is not a problem in large group presentations.
5. Even in the team approach, students tend to identify with one member.
6. The time required to reach consensus in the team planning sessions is often very slow.
7. A common planning period following a large group presentation is better than before.
8. Teachers require training and practice in how to conduct small group seminars.
9. Notebook-keeping and note taking must be taught if large group lectures are used.

BIBLIOGRAPHY

Anderson, Robert H. & Mitchell, Donald P. "Team Teaching—New Learning Concepts Demand Changes in School Plant Design," *The Nations Schools,* 45:75-82, June, 1960.

---

[1] John Schmitt & others. *Problems and Resolution in Design and Execution of Curriculum Research.* Report to the National Association for Research in Science Teaching. Chicago, Feb. 1966.

[2] John Schmitt & others. *Cooperative Development of Unit Achievement Tests in High School Biology.* Report to the National Science Teachers Association. New York City, April 1966.

Bair, Medill & Woodward, Richard. *Team Teaching in Action,* Houghton-Mifflin Co., Boston, 1964.

Beggs, David W. *Decatur-Lakeview High School: A Practical Application of the Trump Plan,* Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

Brownell, John A. & Taylor, Harris A. "Theoretical Perspectives for Teaching Teams," *Phi Delta Kappan,* 43:150-157, Jan. 1962.

Johnson, Robert H., Lobb, D. M., & Swenson, L. G. "An Extensive Study of Team Teaching and Schedule Modification in Jefferson County Colorado School District R-1," *Bulletin of the National Assoc. of Secondary School Principals,* 44:79-93, Jan. 1960.

Johnson, R. H., & Shutes, Robert. "Biology and Team Teaching," *American Biology Teacher,* 24:247-255, April 1962.

Michael, Lloyd S. "Team Teaching," *Bulletin of the National Assoc. of Secondary School Principals,* 47: 36-43, May 1963.

Pella, Milton O. & Poulos, Chris. "A Study of Team Teaching in High School Biology," *Journal of Research in Science Teaching,* 1:232-240, 1963.

Shaplin, Judson T. & Olds, Henry F., Jr., *eds. Team Teaching,* Harper and Row, New York. 1963.

Trump, J. Lloyd & Dorsey, Baynhan. *Focus on Change: Guide to Better Schools,* Rand McNally, Chicago, 1961.

White, Robert W. "The Relative Effectiveness of a Team Teaching Method in High School Biology Instruction," Ph.D. Thesis, University of Wisconsin, 1963. University Microfilms, Inc., Ann Arbor.

PERSONALITY AND ATTITUDINAL CORRELATES OF ALIENATION IN A HIGH SCHOOL POPULATION

James V. Mitchell, Jr. & Roger M. Weir, The University of Rochester

The data to be reported here were gathered in the course of an intensive study of alienated and rebellious students within an academically oriented suburban high school in a major metropolitan area of New York. The school has high expectations for its students, and well over 90 percent of its graduates typically go on to advanced training. The student body is of high socioeconomic status, generally upper middle class or better, and the mean SCAT percentile rank is in the 80-90 range. Like other schools of this type, however, this school, which we shall henceforth refer to as "Suburbia" High School, has its share of alienated students.

In the course of this study the following instruments were administered in this school to the entire class of 1965:

1. The California Psychological Inventory (Gough, 1957)
2. The SRA Youth Inventory (SRA, 1949)
3. The Brown-Holtzman Survey of Study Habits and Attitudes (Psychological Corporation, 1956)
4. A "Personal Satisfactions" Index. This instrument was developed by the authors to gain some insight into the personal values of the respondent. It comprises 16 statements describing different personal satisfactions one might have as an adult, and the subject is instructed to rank these statements in terms of how important he feels that each would be for him personally as an adult.
5. Two student questionnaires. Both of these were developed by the authors, the first to obtain information about aspirations, evaluation of high school experiences, participation in extracurricular activities, etc., and the second to assess the student's reaction to several specific aspects of the high school environment.

Also included in the data analysis were the class ranks for the School and College Ability Test, the class ranks for a cumulative three-year grade point average, and a rough index of socioeconomic status secured from occupational data obtained from one of the questionnaires.

The following procedures were employed in the data analysis. The second "Student Questionnaire" referred to above contained 30 statements frequently heard by Mitchell in the interviews conducted with students known to have attitudes of alienation or rebellion toward the school. Thus this instrument was reflective of the several sources of alienation or rebellion frequently mentioned by the students themselves. Subjects responded to each item of the scale in terms of a Likert-type five-point response continuum indicating varying degrees of agreement or disagreement. As a first step in the analysis the 30 items of this questionnaire were intercorrelated and factor analyzed. The factors emerging from this analysis were interpreted as basic dimensions of alienation within this high school population.

As a second step in the analysis, factor scores were computed for each of the basic dimensions of alienation identified in the first step of the analysis. These factor scores were then correlated with the subscores of the California Psychological Inventory, the SRA Youth Inventory, the Survey of Study Habits and Attitudes, with the items of the "Personal Satisfactions" Index, and with several of the Likert-type items in the first student questionnaire.

Results

The male subgroup had 98 subjects and the female subgroup had 104 subjects. Of the 10 rotated factors that met the eigenvalue criterion of 1.00, 6 of these appeared to be sufficiently stable and meaningful for interpretation. These 6 factors were essentially the same factors for both male and female subgroups, and they were therefore assigned the same labels. Factor 1 is called the "School Rejection" factor, for it has high loadings for all statements expressing dislike of teachers, students, school, and everything associated with it. The second factor has been named "Imputation of Snobbery," because all high-loading items refer to snobbishness in students or a rigid social system. Factor 3 seemed to refer to "Study and Motivational Problems" and was so named. Factor 4 was called "Social Isolation," because items with high loadings referred to feelings of non-belongingness, isolation, and insecurity. Factor 5 was less well determined than the first four factors but had its highest loadings for items expressing a belief that other students were better off financially than the respondent. It is therefore called "Social Status Sensitivity." Factor 6, also less well determined, has high loadings for items expressive of an orientation toward vocational education; it was therefore labeled "Interest in Vocational Education." Certainly the first five factors represent basic dimensions indicative of different kinds and sources of alienation within the high school population, and an "Interest in Vocational Education" within a strongly academically oriented high school such as this one could also serve effectively to alienate the student from the prevailing values of his school culture.

The reader is free to interpret, as he sees fit, any male-female differences in factor pattern. The greatest differences are for the last factor, which is the least well determined for both the sexes. The item referring to parental pressure has a higher loading on "Study and Motivational Problems" for the boys than for the girls, probably reflecting the greater concern on the part of parents for the occupational implications of lack of motivation for boys. The statement referring to ". . . being bothered by the way I act . . ." (item 15) has a higher loading on "Social Isolation" for the girls than the boys, suggesting that slightly different determinants may be at work here in defining the context of social isolation for the two sexes. But the general picture is one of basic similarity between factor patterns.

When the factor scores computed from these factors were correlated with the various subscores and indices of the other instruments employed in the study, the following results were obtained.

1. There is general evidence of consistent negative relationships between desirable and adaptive personality characteristics and the various forms of alienation. It is the students (who are) least (well) equipped in terms of personality characteristics who are most likely to report feelings of alienation of various kinds. This obviously raises two important issues: Is the source of the alienation reported attributable to the environment *per se,* or is it a function of perceptual inaccuracy fostered by personal inadequacies, or is it a resultant interaction between these two? And secondly, are some current attempts to analyze and compare school environments likely to be misleading because of possible perceptual distortions arising from the influence of maladaptive personality traits in some student subgroups?

2. The two subtests in the California Psychological Inventory most related to the various kinds of alienation are Socialization and Achievement via Conformance. Interpretive data provided in the CPI manual suggest that these both reflect the nature of superego development and that both pre-suppose some willingness on the part of the student to *conform* to the requirements and expectations of his environment. Thus the factor of *nonconformity* may be the most critical factor in reported alienation. There is the suggestion in the data that the nonconformist is at least partially guilty of self-inflicted alienation by virtue of his unwillingness to pay the price of conformity. Again we get involved in a "chicken-and-egg" controversy, and again the possibility of interaction effects seems great.

3. The females in the sample generally exhibit higher correlations than the males.

4. Both scores for the School and College Ability Test and the Survey of Study Habits and Attitudes tend to be negatively and significantly related to many varieties of alienation. The same is true to a lesser extent of cumulative grade point average. Poor study habits and attitudes may well go hand in hand with inadequate superego development and unwillingness to conform, and this combination may then be exacerbated by the frustration of a comparatively limited intellectual potential. The self-inflicted alienation referred to earlier may then have fertile ground in which to develop.

5. The SRA "My Home and Family" score, which is the simple total of the number of problems checked by the respondent in this area, tends to have comparatively high correlations with the alienation factors—and often higher than those for the SRA "My School" problem category. Alienation from school seems to be associated with problems at home, and the limitations of behavioral science methodology force us to conjecture: Is it the school problems that cause the problems at home; is it an inadequate home environment that ill prepares the student for school challenges and thus predisposes him toward alienation; or is the alienation a generally pervasive factor, causes unknown, that influences both home and school attitudes alike?

Because of the experimental nature of the scale, the results of the "Personal Satisfactions Index" are admittedly tentative but still suggestive. For the males, School Rejection is likely to be associated with a personal value system that is work-oriented and directed toward the immediate attainment of material goals but less likely to be associated with an emphasis on a good family life and adherence to high moral standards. Study and Motivational problems in boys are associated with interest in an active social life and a "fun" orientation, lack of cultural interests, disinterest in any kind of social contribution, disinterest in adherence to high moral standards, and orientation toward the attainment of material goals. The "life of play" syndrome is apparent here, and it is not surprising to find it correlated with School Rejection. It may also be observed that the work orientation is associated with Social Isolation and Social Status Sensitivity for boys; this suggests the operation of a defense mechanism to permit escape from the unpleasant realities of life into the security of work. For the girls the pattern is not as clear. Lack of intellectual interests is associated with School Rejection, Study and Motivational Problems, and Interest in Vocational Education. Interest in an adult life with an emphasis on hobbies or sports tends to be associated with School Rejection.

Correlations between the various alienation factors and the items of the first student questionnaire show that the alienated tend to find their high school courses less interesting and useful, recognize that they have study and motivational problems, and acknowledge that their parents are not satisfied with their grades and that they have not lived up to parental expectations.

The following generalizations seem to be justified from the results of this study. There is evidence of several varieties of alienation within this high school population. These are related to maladaptive personality characteristics, which makes it difficult to determine whether the causal origins of the alienation are to be located primarily within the environment, within the individual, or in the interaction between the two. Most closely related to the various forms of alienation are those personality characteristics involving an unwillingness to conform to the expectations of one's social milieu. The alienated can also be discriminated from others in terms of their lower scholastic aptitude, their poorer study habits and attitudes, their value orientations, and their greater incidence of reported home problems. The existence of a personality and attitudinal syndrome associated with alienation is clearly apparent, and especially so with the factor of School Rejection, but the causal relationships involved here are admittedly complex and will require much research for their final explication.

VALIDITY CONCOMITANTS OF CUMMINGS WORLD HISTORY TEST, OTIS QUICK-SCORING MENTAL ABILITY TESTS: GAMMA TEST, AND NEW YORK STATE REGENTS EXAMINATION IN WORLD HISTORY

Norman C. Maberly & Harold F. Bligh, Harcourt, Brace & World, Inc.

Teachers preparing students for the New York State Regents Examinations should be interested in information that will reveal weaknesses in their instructional programs. Among the more useful types of information for this purpose is that which will enable them to predict, to some extent, the potential for success of the pupils, and the particular areas in which they may be relatively strong or weak. The use of such data presupposes, of course, that empirical evidence is available attesting to its predictive validity.

Purposes of the Study

Since an interest has been shown in using published standardized tests as predictors of performance on the New York State Regents Examinations, the Test Department of Harcourt, Brace & World, sought the cooperation of two New York State high schools in a study investigating the interrelationships of three variables in the area of world history as follows:

Cummings World History Test, Form E, 1966, raw score
Otis Quick-Scoring Mental Ability Tests: Gamma Test, Form Fm, IQ
Board of Regents Examination in World History, 1965, raw score.

Through an analysis of these variables, correlation coefficients and regression equations were derived which permit a comparative evaluation of the predictive validity of the standardized tests when the criterion is the Regents Examination in world history.

The Sample

The sample consisted of tenth-grade students enrolled in world history classes in two New York State public high schools during the spring of 1965. These two schools were members of a nationwide population participating in the standardization of the revised edition of the Cummings World History Test. The students within each school were divided randomly into two groups. The first group took Form E of the Cummings World History Test, the second, Form F. Both groups took the Otis Quick-Scoring Mental Ability Test. The findings reported here are for the group taking the Form E test.

Procedure

The data were collected in two steps. In the first phase, participating students were administered the Cummings test and the Otis test concurrently. Several weeks later they took the Regents test as part of the regular New York State Regents Program. The data were analyzed separately for males and females, according to the course of study being pursued. Finally, the data were combined to give total group statistics.

Results and Discussion

Table I summarizes the data for the two schools combined. Of particular interest is the fact that the boys appear to perform a little better than the girls in world history as demonstrated by scores on both the Cummings test and the Regents examination. This difference persists even where the girls are significantly better than the boys in mental ability.

A consideration of the correlation coefficients reveals that the Cummings test has a high relationship to the Regents Examination, especially for students in college preparatory groups. For the combined sexes, the correlations are .72, .66, and .76 for the total group, general, and college preparatory groups respectively. The Otis IQ, however, does not show such a high relationship as evidenced by correlations of .50, .40, and .55 for the same groups.

The influence that mental ability components may have upon the relationship of the other two variables, is illustrated through an analysis of the residuals when mental ability is held constant. "The partial correlation $r_{13.2}$ describes the correlation remaining between the Cummings test and the Regents Examination when the IQ influences are statistically held constant. For the total group of 262 pupils, $r_{13.2} = .61$."

From these data it would appear that the Cummings World History Test, Revised Edition, has potential as a predictor of success for students aspiring to the Regents Examinations in world history, and that by prior administration of the test, teachers may obtain useful information with regard to the possible success ratio for their individual students and groups.

To further facilitate interpretations of the Cummings data, the following regression equations were derived:

Total Group Regents Score = .90 (Cummings Score) + 39.

General, Etc. Regents Score = .96 (Cummings Score) + 29.

College Prep. Regents Score = .97 (Cummings Score) + 27.

In the use of these equations, cognizance must be taken, of course, of the dependence of regression data on particular groups. These equations were based on data accumulated for two schools in New York State having specified characteristics. The generality of the equations to other schools and situations has not been established.

In interpreting the results presented here, it is important to remember that the sample of students includes only those who took both the Cummings test and the Regents. Students working for school credit only, or who were absent on the day that the Cummings test was administered were excluded. Thus the correlations tend to be conservative estimates of the existing relationships.

A comparison of the Cummings results with those for the standardization groups reported in the Cummings *Manual* offers further information on the nature of the sample. In Table II, it can be seen that the mean Cummings raw scores for the Regents groups run about 12 points higher than those for the corresponding norms groups; the standard deviations run from three to four points lower, suggesting a restriction in range of raw scores for the Regents groups.

TABLE I

CUMMINGS WORLD HISTORY TEST, FORM E, and OTIS QUICK-SCORING MENTAL ABILITY TESTS: GAMMA TEST: Intercorrelations with New York State Regents Examination in World History.

		1. CUMMINGS		2. OTIS IQ		3. REGENTS		Correlations		
	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	12	13	23
Total Group										
Male	128	48.3	11.0	112.1	10.2	74.3	13.1	.64	.68	.49
Female	134	46.2	9.5	114.4	9.0	73.2	13.5	.51	.77	.54
Combined Sexes	262	47.2	10.3	113.3	9.7	73.7	13.3	.56	.72	.50
General, Etc.[1]										
Male	52	43.7	10.1	107.2	9.3	72.7	12.2	.52	.62	.51
Female	66	43.0	8.8	111.6	8.5	69.7	14.5	.32	.70	.39
Combined Sexes	118	43.3	9.3	109.6	9.1	71.0	13.6	.40	.66	.40
College Prep.										
Male	76	51.5	10.4	115.5	9.4	75.4	13.6	.63	.74	.48
Female	68	49.3	9.3	117.2	8.7	76.5	11.6	.57	.82	.64
Combined Sexes	144	50.4	9.9	116.3	9.1	76.0	12.7	.59	.76	.55

[1] General, Etc., includes all students who did not specifically indicate that they were following a college preparatory program in high school.

TABLE II

CUMMINGS WORLD HISTORY TEST, FORM E:

Comparison of the Regents Samples and the Standardization Norm Groups.

		REGENTS SAMPLE				FORM E NORM GROUP		
	N	Mean	S.D.	%ile Rank	Stanine	N	Mean	S.D.
Total Group								
Cummings Raw Score	262	47.2	10.3	81	7	2,325	34.2	13.7
Otis Gamma IQ	262	113.3	9.7	66	6	2,325	106.3	13.7
General, Etc.[1]								
Cummings Raw Score	118	43.3	9.3	80	7	1,305	31.6	12.4
Otis Gamma IQ	118	109.6	9.1	66	6	1,305	104.4	11.7
College Prep.								
Cummings Raw Score	144	50.4	9.9	80	7	1,010	37.6	14.5
Otis Gamma IQ	144	116.3	9.1	62	6	1,010	108.8	15.6

[1] General, Etc., includes all students who did not specifically indicate that they were following a college preparatory program in high school.

STUDENT DRINKING HABITS AS RELATED TO PERSONALITY VARIABLES

S. W. Johnson, Thomas Carlsen, & M. L. Levine, State University College at Plattsburgh

This study reports the findings gathered from a study of 100 college students who participated in experimental examination of field dependency and subsequent analysis of the etiology of their drinking habits.

Subjects were tested for field dependency responses on the Witkin RF Test and then responded to a detailed questionnaire regarding how and when they formed their drinking habits, their feelings regarding various drinking situations, and self and projected attitudes regarding these habits and situations.

It was found that female subjects appear to have a more field oriented gestalt regarding the whole alcohol consumption pattern than males and that females in the college sample also appear to show strong generalized guilt feelings about their own habits and attitudes relating to liquor related behavior. Such feelings are evidenced in reported self-deprecating ratings of their own self-control and drinking behavior pattern as compared to peers; higher concern with negative social aspects of their parents' drinking behavior, and significantly higher ego-ideal behavior pattern where alcohol in concerned.

Extensive information regarding general alcohol consumption habits and background for such habits is also reported.

STUDENTS' PERCEPTIONS OF COLLEGE EXPERIENCES RELATED TO DROPOUT AND ACHIEVEMENT

Horace B. Reed, Skidmore College

This research focuses on college student motivations related to criteria of voluntary dropout and of under—overachievement. On the basis of leads from an eclectic selection of psychological, sociological, and educational generalizations, motivational constructs are described, and hypotheses made, concerning relationships to the two criteria.

A 56 item College Assessment Inventory was developed to measure the antecedent motivational constructs, with each assessed by four to ten items. Student responses to the five point scaled items measure the students motivations within the context of varied college experiences: classes, readings, papers, evaluations, instructors, administrators, peers, extracurricular and social life. Evidence for the reliability of the Inventory items is furnished by the close similarity in the distribution of responses for samples of a freshman class and of a sophomore class from the same college. Validity of the items can be partially inferred from their predictive powers.

The 97% sample includes 343 college freshmen of the 1966 graduating class from a private, resident women's college offering liberal arts and professional curricula. With the intention of distinguishing among types of dropouts, voluntary attrition is defined as any student who does not graduate with his class, and who leaves for other than such compulsory reasons as disqualification, illness, and sickness at home. The under-overachievement criterion is defined as the difference between the student's predicted grade point average (based on Scholastic Aptitude Test scores and High School Rank) and the freshman year grade point average; this criterion has a built-in control for differences in academic ability and previous achievement.

Hypotheses for relationships between the motivational items and the voluntary dropout criterion are tested with chi square; analysis of variance is used with the under-overachievement criterion. Acceptance of a .05 significance level or better is selected for both tests.

The motivational variable of *meaningfulness of daily college tasks* (the student's assessment of the relevance of daily academic work to his present interests and concerns) is positively related to overachievement; seven items demonstrate that students perceiving relatively high intrinsic meaning in their work tend to overachieve. None of the *meaningfulness* items are related to the attrition criterion.

On the antecedent variable of *relevance of college to the student's future goals* (the student's long range goals and his evaluation of present college attendance as an effective means of attainment), higher scores on long range goals are less predictive of overachievement, with only one item reaching significance. With the voluntary dropout criterion, three items indicated that high scores on *future goals* are positively related to staying in college.

Warmth of interpersonal relations (the student's feelings concerning the warmth of interpersonal relations within the college community) is not related to under-overachievement; with the attrition criterion, three items indicate that high scores on *warmth* are positively related to graduating.

Scattered evidence in predicting one or the other of the criteria is demonstrated for the motivational variables of academic challenge and of importance of grades. No evidence is found to support the predictive powers of two other variables: the college's concern for the individual, and the college's implementation of selected educational goals. Preliminary detailed analysis of items for these four variables suggests that in some instances rewording refinements, and more relevant selection of college experience contexts, may be worth exploring before rejecting these variables.

Another type of motivational measure, called *field of interest* (the student's present preference for a college major or area), was also included in the research design. Significant differences in mean under-overachievement scores, and in percentages of voluntary and of total dropouts, occur among the twenty-one fields of interest. Comparing the liberal arts (14 fields) with the professional (7 fields), the mean under-overachievement scores are similar. The percentage of voluntary dropouts, and of all dropouts, was significantly larger in the liberal arts group of fields than in the professional group.

In conclusion it was demonstrated that several of the College Assessment Inventory motivational items are useful in predicting a significant amount of the variation in the college under-overachievement scores, and in distinguishing between potential voluntary college dropouts and those who graduate. The item content may provide practical clues for increasing achievement or reducing attrition through counseling, and through modifications in college programs and teaching procedures.

COOPER UNION OAIS VALIDITY STUDY—1966

Walter S. Watson, The Cooper Union

The **Opinion Attitude And Interest Survey** (OAIS) is a personality test which was designed by Benno Fricke at the University of Michigan. In two earlier reports to this association, I noted that OAIS scores were not readily subject to faking and that attempts to use the OAIS to increase the accuracy of freshman grade predictions showed confusing results.

At the top of Table I, you will note that Group A entered in 1962, Group B in 1963 and Group C is a combination of all engineering and science students entering in 1962 or 1963 who completed at least two years of study. It includes six students who entered in 1962 and completed two years, but who did not finish the three years needed to be counted in Group A. The somewhat cryptic column headings for Table I or Table III of "Usual Predictors" or "OAIS Scores" can be understood by reference to the factor items in the following Tables of Standard Regression Weights. We found a multiple correlation of .5187 between freshman engineering grades and our usual predictors. When 3 OAIS scores were added, the correlation with first year grades was .5794. The difference between .5187 squared and .5794 squared represents an increase of 24.8 percent in the predictable variance of freshman grades for the 1962 entering class. Predictions of the variance for their sophomore and junior grades were improved by 30.3 percent and 34.9 percent respectively when we added the 3 OAIS scores.

However, when our usual predictors hit the target with a multiple r. of .6929 as they do for the 1963 class, we get only a 4.6 percent increase in predictable variance, and the combined classes in Group C show a low improvement of 10 percent or less. If we look at the standard regression weights for Group C in Table II, we can conclude that there were 3 or 4 areas worth using. The old standby, "high school average," leads the parade. Advanced Mathematics comes second and Physics stands third for predicting freshman grades, but retreats to fourth rank for sophomore grades when the OAISs scores for Achiever Personality are included as predictor variables.

Tables III and IV offer similar data for groups entering The Cooper Union School of Art and Architecture in 1962 and 1963. Two of the OAIS scores which we used to improve our art and architecture predictions are the same as those used for engineers, Creative Personality and Achiever Personality. The third OAIS score is a measure of interests like those who major in biology (Bio Int). Why such interests are related to Cooper Union art and architecture grades is unexplained, but apparently they are related. We find that the lower the correlation with our usual predictors, the greater the increase when we add the OAIS scores. When the usual predictors approach their customary peak, in this case .5851 for Group E, year 1, then the proposed new tests add little. When the usual tests go haywire as they did with an r. of .2324 for Group D, year 1, then appropriate weighting of the OAIS tests can add substantially to the predicted variance.

Inspection of the regression weights for the combined groups shows the old standby "high school average" is still a strong, steady predictor of college grades. The Cooper Union test of design aptitude as a predictor of second year grades at Cooper Union shows only a low positive weight for freshman grades, but bounces back to its usual equality with high school grades.

The reader is warned that the limited range of the engineering and art populations at Cooper Union contribute to any study aimed at the average grades of our rather pure-vocational samples, but when anyone attempts to predict differences within these restricted ranges of design or engineering aptitude, the variability may be unrelated to parameters used in selection for admission. I gave the OAIS again in 1966 and propose to follow these freshmen for at least two years as a final check. In spite of my doubtful validity findings, I do not know of any other more promising personality measure designed for mass use in the college admission and course selection process.

TABLE I
ENGINEERING AND SCIENCE CLASSES ENTERING IN 1962 AND 1963
GROUP A-1962 N 86, GROUP B-1963 N 109, GROUP C-COMBINED N 201
MULTIPLE CORRELATIONS

HPR Criterion	Usual Predictors	Usual Predictors +OAIS Scores	Percent Increase In Mult Corr
	GROUP A		
Year 1	.5187	.5794	24.8
2	.5322	.6076	30.3
3	.4687	.5443	34.9
	GROUP B		
Year 1	.6929	.7087	4.6
2	.6357	.6644	9.2
	GROUP C		
Year 1	.5966	.6154	6.4
2	.5373	.5642	10.3

TABLE II
REGRESSION WEIGHTS FOR GROUP C

Factor	Usual Yr-1	+OAIS	Usual Yr-2	+OAIS
HS Aver	.3275	.3002	.3137	.2745
SAT-V	−.0171	−.0312	−.0289	−.0009
SAT-M	−.0099	.0193	.0254	.0612
Chem Ach	.0501	.0429	.0765	.0546
Phys Ach	.1957	.1666	.1494	.1316
Adv Ma	.2415	.2456	.1764	.1694
OAIS				
Ach P	....	.1499	....	.1630
Cr P	....	.0686	....	−.0741
Em Adj	....	−.0192	....	−.0301
Mult Corr	(.5966)	(.6154)	(.5373)	(.5642)

TABLE III

ART AND ARCHITECTURE CLASSES ENTERING IN 1962 AND 1963

GROUP D-1962 N 63, GROUP E-1963 N 52, GROUP F-COMBINED N 140

MULTIPLE CORRELATIONS

HPR Criterion	Usual Predictors	Usual Predictors + 3 OAIS Scores	Percent Increase In Mult Corr
GROUP D			
Year 1	.2324	.4715	311.6
2	.2947	.4350	117.9
3	.3261	.4404	82.4
GROUP E			
Year 1	.5851	.5968	4.0
2	.4593	.4891	15.7
GROUP F			
Year 1	.2229	.3253	113.0
2	.4139	.4670	27.3

TABLE IV

REGRESSION WEIGHTS FOR GROUP F

Factor	Usual Yr-1	+OAIS	Usual Yr-2	+OAIS
HS Aver	.1979	.1455	.3210	.2908
SAT-V	.0773	.1167	−.0163	−.0056
SAT-M	−.0477	−.0309	−.0306	−.0053
Design Apt	.0779	.0622	.3161	.2968
OAIS				
Ach P	....	.1406	....	.0683
Cr P	....	−.0045	....	.1451
Bio Int	....	.1571	....	.1807
Mult Corr	(.2229)	(.3253)	(.4139)	(.4670)

FOUR YEAR DEVELOPMENT OF COLLEGE STUDENTS' VALUES

S. W. Johnson, State University College at Plattsburgh

This study reports progressive value changes by students of each sex in five major curricula during four years of college. (Sample size approximated a minimum of 50% or more of all students enrolled in that curricula.)

Students were requested to respond to the Poe Inventory of Values during the last month of the spring semester of each of their four collegiate years. Value measurements were obtained in eight areas (aesthetic, intellectual, material, power, social contact, religious, prestige, humanitarian) by use of this inventory.

Analysis of the collected data indicates considerably more value change than previously reported in other longitudinal value studies. This may be due to the design change which has enabled progressive study rather than a simple test-retest analysis.

A general regression toward group means following some significant change is apparent in many areas. Such a regression would cancel out test-retest differences. Yet it would yield statistically significant results when submitted to analysis of variance considering four different points between initial and final testing. Thus, changes do occur but are sometimes cancelled out by group regression trends.

Of the eight curriculum-sex groups studied, General Elementary Education female students showed the most changes with significant shifts on 6 of 8 value areas during the four years. Liberal Arts males showed the least change with only one value area showing significance of shift for the same period.

Examined by value areas, social contact values appeared to be least stable with all eight subgroups showing significant decrements in the strength of values in this area over four years. Power and Aesthetic values showed no change in any group, while Humanitarian and Material value areas changed for all but one of the eight subgroups.

A total of 19 different significant shifts of a value area by a subgroup were recorded. Thus nearly one third of the possible shift areas did move to a significant degree. Thirteen of the nineteen significant findings which occurred, seemed to have been influenced largely by shifts during the junior year. Five were predominantly influenced by sophomore year changes and one showed a steady change from the sophomore year on.

Follow up studies of graduate students from this population are planned.

SOCIAL DISTANCE AND DEVIANCY AMONG COLLEGE STUDENTS

Louise W. Fox, Finch College & David J. Fox, The City College

Statement of the Problem

This investigation replicated Bogardus' 1928 study of social distance but with social, physical and political groups and ideas rather than racial and ethnic groups. Subjects were asked to indicate the maximum degree of acceptance they felt toward others deviant from the average in a particular political idea, physical characteristic, or social behavior.

The Questionnaire

A questionnaire was devised employing Bogardus' scale of Social Distance. The respondents were asked to indicate how they felt about 20 immigrants applying for permission to enter the United States, using Bogardus' seven-interval scale reflecting different degrees of social distance (or acceptance). These (listed on Table I) range from, "I would exclude this person from the U.S.A.", to "I would admit this person to my family through marriage".

However, where Bogardus had his subjects respond to racial or ethnic groups, this study used ten social-physical and ten political deviants (also listed on Table I). For example, included in the social-physical deviants were a person of normal intelligence but illiterate, a person totally blind, a prostitute, and an alcoholic. Included in the ten political deviants were persons who believed in the philosophy of Franco, the Russian Communist Party, Swedish or British Socialism and the John Birch Society.

The instructions given to each subject were:

> In this survey you are asked to assume that twenty individuals have applied for admission to the United States as immigrants. Each of the twenty has a characteristic that has made the immigration officers hesitate to admit them. In an effort to make policy on this issue, you and others like you have been asked to indicate the *maximum degree* of closeness you would prefer to establish with each of the twenty individuals, on the basis of the information given.

Subjects

The instrument was administered the same day to the 50 female and 28 male members of three different sections of the same course in Education. The students were not asked to identify themselves by name, but only to state sex, age, major subject area in school and political affiliation if any. For males the median age was 21.4 years, for females 19.9 years.

Methods of Data Analysis

For purposes of this study the categories indicating that the S would accept the potential immigrant into his family through marriage or to close friendship were combined and considered indicative of *complete acceptance.* Similarly the categories indicating that the S would exclude the potential immigrant from the U.S.A. or would permit him to enter as a visitor only were combined and considered to indicate *rejection.* The three intermediate categories were considered to indicate *limited acceptance.*

Results

On the average, the S's completely accepted 4.6 deviants. In comparison, they rejected an average of 6.7 deviants, and the difference between these means was statistically significant at the .01 level.

The range in number of the deviant groups completely accepted was between 0 and 13. Twelve (15.4%) of the S's completely accepted no one and twenty-two (28.2%) completely accepted only one, two, or three deviants and only eight of the 78 completely accepted as many as half of the deviants.

Considering rejection, the range was from the eight S's who rejected no one, to the one S who rejected 16 of the 20 deviants. Twenty-one (27 percent) of the S's rejected half or more of the deviants and an additional 30 (38 percent) rejected between five (the mean number *accepted*) and nine.

Comparing the acceptance of social-physical and political deviants, there were no significant differences in terms of complete acceptance, however significantly more social-physical than political deviants received limited acceptance and significantly fewer were rejected. Moreover, 43 S's completely accepted more social than political deviations, compared to the 11 who completely accepted more political than social deviants.

[Table I presents for males and females] Now consider, by sex, the percent of complete acceptance, limited acceptance, and rejection for each stimulus group.

The table lists the deviants in rank order for complete acceptance for the total groups studied. The most completely accepted deviant was the blind person: 71.7% of the total group were willing to accept the blind person into kinship through marriage or as a close personal friend. Similarly, the British Socialist, the nonhospitalized person in psychoanalysis, and the Swedish Socialist were completely accepted by more than half of the subjects. About ⅓ of the subjects were completely accepting of the illiterate and hospitalized mentally ill persons.

Of these six most accepted groups, two, the British and Swedish Socialist, were of a political nature, while the remaining four were social-physical deviates.

The other 14 deviants were completely accepted by only small proportions of the S's, with the smallest proportions accepting the Apartheidists, Chinese Communists and heroin addicts. As might be expected, those least often completely accepted were most often rejected, and vice versa. A rank order correlation of —.94 showed the magnitude of this inverse relationship to be very high.

Several distinct groups emerge on the basis of complete acceptance and rejection of the S's by sex. The three groups accepted by significantly larger proportions of women than men were all social-physical and the three groups accepted by significantly larger proportions of men than women were all political.

Conversely all four groups rejected more frequently by females than males were political and the three groups the men rejected more frequently than the females were social-physical.

Discussion and Conclusions

The hypothesis that there would be a greater amount of acceptance than rejection of the deviant was not supported, and in fact, was contradicted. The average S rejected significantly more deviants than he accepted. If this tendency to reject deviation is considered evidence of conformity, then this finding would suggest that the college students studied were relatively conforming.

The second hypothesis, that there would be greater acceptance of the social-physical deviants than of the political deviants was supported. This finding may be related to several different factors. First, many of the social-physical deviants, *i.e.,* the blind, illiterate, the retarded, were probably nonthreatening to the normal person. In contrast, the political philosophies may well have been threatening, particularly in view of that amorphous yet powerful force we call "political climate."

A second possible explanation of the greater degree of acceptance of social-physical rather than political groups may be the factor of individual responsibility for deviation, *i.e.* the blind and mentally retarded may be perceived as victims of a condition over which they have no control, and for which they should not be rejected. On the other hand respondents may have felt that since an individual in our society can select his own political philosophy he must carry responsibility for his choice.

The major implication of these findings for the researcher is that in view of the overall pattern of rejection rather than acceptance, the traditional view of the college student as a liberal, freethinking, experimentalist, needs reevaluation. If actually this is a more *liberal* group, then the question arises as to what is the range of acceptance of deviation among the noncollege population, *i.e.,* the less *liberal?* If it is even less, how does this reflect on the democratic process and the ability of the society to grow through the absorption of new and different ideas? If, on the other hand, the noncollege population is *more* accepting, what are the factors in operation which produce this lack of acceptance of deviation among college students? How does the educational process affect these attitudes? These issues, while not under consideration in this study, seem to the researcher to be worthy of further investigation.

TABLE I

For Each Stimulus Group the Per Cent of Complete Acceptance, Limited Acceptance and Rejection by Sex of Respondents

Stimulus Group	Per Cent Complete Acceptance[1]		Limited Acceptance[2]		Per Cent Rejection[3]	
	Females	Males	Females	Males	Females[1]	Males[2]
1) Blind	72	71	28	21	00	08
2) British Socialist	60	65	30	28	10	07
3) Non-Hospitalized Psychiatric	72	35	26	57	02	08
4) Swedish Socialist	60	65	30	28	10	07
5) Illiterate	30	46	64	47	06	07
6) Hospitalized for Mental Illness	38	22	46	56	16	22
7) Non-Self Supporting	16	21	50	51	34	28
8) French Algerian	08	29	66	53	26	18
9) Castroite	12	18	24	24	64	58
10) Mentally Retarded	14	11	64	66	12	25
11) John Birchite	06	22	38	53	56	25
12) Alcoholic	14	07	64	75	22	18
13) Homo-Sexual	16	04	66	50	18	46
14) Russian Communist	14	07	20	33	66	60
15) Prostitute	12	07	58	58	30	35
16) Francoite	10	11	36	46	54	43
17) Heroin Addict	10	04	46	36	44	60
18) Chinese Communist	02	11	17	21	81	68
19) Apartheidist	02	11	36	53	62	36
20) Nazi	00	07	10	22	90	71

[1] Complete Acceptance—1. I would admit this person to my family through marriage
2. I would admit this person to close personal friendship

[2] Limited Acceptance—3. I would admit this person to my neighborhood to live
4. I would admit this person to my major field as a fellow worker
5. I would admit this person to the U.S.A. as a citizen

[3] Rejection —6. I would admit this person to the U.S.A. only as a visitor
7. I would exclude this person from the U.S.A.

CRITICAL REVIEW OF THE STANFORD-BINET IQ TABLES

Conwell Higgins, Albany Public Schools

With respect to IQ variability in relation to age, Terman and Merrill (1960, p. 16) explained:

> If the scale is to yield comparable IQs at all age levels, it is essential that the standard deviations at the various levels shall be approximately equal. The use of IQ rests on the assumption that variability in terms of IQ remains approximately constant from age to age; in other words, *that variability in terms of mental age is directly proportional to chronological age. (Italics added.)* (Only to the extent that this assumption is true does the IQ have constant meaning.)

A clear distinction is made between defects of IQ Tables and defects of any particular age scale. An IQ Table may meet the criteria of a linear age scale or it may reflect the nonlinear unit defects of any particular age scale.

The question of IQ Table defects rests upon the answer to this question: Is mental age variability directly proportional to chronological age? It will be demonstrated, that for ages 2-16, (1) the mental age variability, of the 1937 SB ratio IQ Tables, is directly proportional to age; (2) whereas the mental age variability, of the 1960 SB deviation IQ Tables, fluctuates over age.

Proportionality of Mental Age Sigmas to Chronological Age

The variability in terms of mental age to chronological age, derived from the SB 1937 ratio IQ Tables and the 1960 deviation IQ Tables is presented in Table I. Row 1 of Table I reads, "At age 2, the 1937 mental age sigma is 3.84 months." Row 2 of Table I reads "For all ages, 2-16, the ratio of 1937 mental age sigmas to chronological age is a constant and uniform value of 0.160." When the ratio value is multiplied by 100, the product is an IQ standard deviation of 16 points, for all ages 2-16. This is incontrovertible evidence that the 1937 IQ Table standard deviations are constant and uniform throughout the 1937 IQ Tables.

The data for the 1960 deviation IQs behave quite differently. Row 3 of Table I reads, "At age 2, the 1960 mental age sigma value is 3.88." Row 4 of Table I reveals that the ratios of 1960 mental age sigmas to chronological age, for ages 2-16, are unstable and fluctuate from a low value of 0.143 at age 6 to a high value of 0.181 at age 12. Thus, using the explicitly stated criterion of Terman and Merrill (1960, p. 16), the mental age variability is not proportional to age but fluctuates so that the 1960 deviation IQs yield comparable IQs from one discrete age to another discrete age only through the deformation of the mental age values.

Proportionality of Mental Age Increments Over Age by z Score IQs

The mental age increments over chronological age by z score IQs, derived from the SB 1937 and 1960 IQ Tables are presented in Table II. Row 1 of Table II reads, "For the 1937 SB IQ of 36, ages 2-13, the mental age score yearly increment is a uniform 4.32 mental age units." Row 2 of Table II reads, "For the 1937 SB IQ of 36, ages 13-16, the mental age score increment is a uniform 2.88 mental age units." Row 3 of Table II reads, "For all the 1937 SB z score integer IQs, ages 2-13, the ratio of mental age increment to z score IQ is a constant value, .12." Row 4 of Table II reads "For all the 1937 SB integer z scores IQs, ages 13-16, the ratio of mental age increment to z score IQ is a constant value, 0.8." Thus, it has been demonstrated that the SB mental age units of the 1937 ratio IQ tables, are linear because the ratios of mental age increment to z score IQs are constant values.

Mental age scores, when used to enter the 1960 deviation IQ Tables, take on sharply different characteristics. Row 5 of Table II reads, "For the 1960 SB deviation score of 36, the mean mental age increment, ages 2-18, is 3.74 mental age months." Row 6 of Table II reads, "For the 1960 SB deviation IQ of 36, the sigma of the mental age increment, ages 2-18, is 2.59 mental age months." Row 7 of Table II reads that, for the 1960 SB z score IQs, ages 2-18, the ratio of mean mental age increments to integer z score IQs fluctuate from a low value of .10 to a high value of .19. Hence, increments of mental age scores derived from the 1960 deviation IQ Tables, fluctuate erratically overage for z score IQs and specify peculiar patterns of intellectual development. For example, the 1960 deviation IQ of 36 specifies no increment in mental age value from age 10 to age 12, a 24 month period. It does however specify an increment of 10 mental age units from age 15 to age 16.

Choice of IQ Tables

In the analysis of the 1937 standardization data, Terman and Merrill (1937, pp. 40-42), found that IQ variability, subsample to subsample, based upon ratios of mental age sigmas to chronological age, was erratic overage. A decision had to be made with respect to IQ Tables. The choices were: (1) Consider the sample as a whole, ages 2-16, and assume that true IQ variability over developmental age is constant; (2) Consider that the observed IQ variability actually reflects the true population variability at each age, and treat each age subsample discretely.

The 1937 decision was to assume that IQ variability overage is constant. Therefore, the ratio IQ Tables were presented. The linearity of the mental age units was maintained; the variability in terms of mental age to chronological age was constant; IQ variability over developmental age was uniform, and IQs were ratio as well as standard scores. Mental age tables, derived from the 1937 IQ Tables, are a lovely lattice and a generalized model of intellectual development. If the mental age sigmas of the 1937 standardization data are normalized to yield the constant ratio MA sigma/CA of 0.16, the mental age sigmas are identical with the mental age sigmas which yield 1937 IQ standard deviations of 16 points. For example, at age 6, observed IQ deviation $12.5_{L\ IQ}$:

IQ standard deviation $16.0_{L\ IQ}$ = observed sigma $9.0_{L\ MA}$:

normalized sigma $11.52_{L\ MA}$.

However, in the preparation of the 1960 IQ Tables, the 1937 decision was reversed. It was assumed that fluctuating IQ variability overage actually reflect true differences in IQ variability overage. By presenting deviation scores based upon the discrete subsample IQ standard deviations of the standardization data, the rationale of linear mental age units and ratio IQs was abandoned.

Thus, if true variability is approximately constant over developmental age, the 1960 deviation Tables are adjusted to reflect sampling errors and artifacts of the Stanford-Binet Tests of intelligence.

Recommendation

It is recommended that the MA scores obtained in intellectual assessment by means of the 1960 Stanford-Binet Intelligence Scale be converted to IQs through use of the 1937 ratio IQ Tables.

BIBLIOGRAPHY

Terman, L. M. and Merrill, Maud A. Measuring Intelligence. Boston. Houghton Mifflin. 1937.

________. Stanford-Binet Intelligence Scale Manual. Boston. Houghton Mifflin. 1960.

THE RELATION BETWEEN A DISGUISED AND REPORTED MEASURE OF RISK TAKING ON OBJECTIVE EXAMINATIONS

Malcolm J. Slakter, State University of New York at Buffalo

Consider a choice-situation in which an individual is faced with a set of actions, each of which is associated with a stake to be lost if the action taken is unsuccessful, and a prize to be won if the action taken is successful. By some decision process, the individual must select a single action from the entire set of possible actions.

One variable conjectured as entering into this decision process may be referred to, for want of a better name, as risk taking. The casual observer will recall that in various choice-situations, individuals appear to differ with respect to risk taking propensities; e.g., games of chance, driving on the thruways, vocational choice, etc.

If a study of risk taking is to be valid, however, the prizes and stakes must be apparent and meaningful to the subjects. Hence, we find that a convenient and promising way to study risk taking is afforded by the ordinary objective examination which warns the examinee that incorrect responses will be penalized. The concern of this paper, therefore, is with the measurement of risk taking on objective examinations.

An objective examination is defined here as one which is composed of items to which the examinee responds by selecting an alternative from those presented in the item. Guessing occurs when an examinee responds to one of these items by randomly selecting the alternative; risk taking is then defined as guessing when the examinee is aware that there is a penalty for incorrect responses.

A disguised measure of risk taking developed by Ziller [2] is defined as

$$R_z = \frac{\frac{n}{n-1}w}{\frac{n}{n-1}w + u},$$

where n is the numer of alternatives for each item,

w is the number of incorrect responses for the examinee, and

u is the number of items unattempted by the examinee.

R_z is an estimate of the proportion of unknown items on the test to which the examinee responded; i.e., the ratio of true number of questions guessed to the true number of questions not known.

For this particular study, the examination used was the Concept Mastery Test (CMT) [1]. The directions of the CMT instruct students to "Omit those items that you could answer only by pure guess, but answer all you *think* you know, even if you are not quite certain." The first part of the CMT has two alternatives and is based upon vocabulary; the second part of the CMT has three alternatives and is based upon verbal reasoning. A slight modification is therefore necessary in the calculation of R_z to allow for the different values of n for the two parts of the test.

Immediately following the administration of the CMT, examinees were asked to complete a questionnaire; one item of which stated:

> CONSIDER THOSE ITEMS ON THE TEST THAT YOU COULD ANSWER BY PURE GUESS ONLY. APPROXIMATELY WHAT PERCENTAGE OF THESE ITEMS DID YOU ANSWER?

Responses to this item were considered to be the reported measure of risk taking (R_z'). The study was replicated in five different classes in introductory educational psychology; three classes at the University of California at Berkeley in the summer of 1965, and two classes at the SUNYAB during the spring semester of 1966. As another facet of the study, it was decided to investigate the ambiguity of the directions. Therefore, another item on the questionnaire stated:

> WHILE TAKING THE CMT, DID YOU ASSUME THAT YOUR SCORE WOULD BE DECREASED FOR EACH INCORRECT ANSWER; I.E., YOU WOULD BE PENALIZED FOR GUESSING?
>
> A. YES
> B. NO
> C. UNSURE

In this report, only the data of students responding A or C will be analyzed; *i.e.*, if the student assumed that there was no stake to be lost, he was dropped from the analysis. A complete report would show that more than 25 percent of the students assumed that there was no penalty for incorrect responses, but the overall characteristics of the disguised and reported measures were essentially unchanged.

Table I presents the number of students (n) in the class that were at least aware of the possibility of a penalty for incorrect responses, the mean and standard deviation for the disguised measure of risk taking $(\bar{X}_Z, S_Z)$ and the mean and standard deviation for the reported measure of risk taking $(\bar{X}_{Z'}, S_{Z'})$.

TABLE I

Means and standard deviations for disguised and reported measures of risk-taking (Values with primes are reported measure; C_1, C_2, C_3 represent Berkeley classes, B_1, B_2 represent Buffalo classes)

	C_1	C_2	C_3	B_1	B_2
n	19	30	9	39	23
$\bar{X}_z$	.77	.69	.74	.71	.50
$\bar{X}_{z'}$	.33	.26	.36	.27	.14
S_z	.22	.20	.22	.21	.20
$S_{z'}$	.44	.32	.32	.34	.12

It is seen that the classes are fairly similar with respect to mean disguised risk taking and also with regard to mean reported risk taking, but there are clear differences between mean disguised and reported risk taking scores within each class. In each case, the mean disguised risk taking score is higher than the mean reported risk taking score. In similar fashion, the classes are quite uniform with respect to the variance of the disguised measure. However, the classes are less uniform with respect to the variance of the reported measure; the reported measure also tends to be more variable than the disguised measure.

Correlations between the disguised and reported measures for each class are presented in Table II. An approximate test of the equality of population correlation coefficients was carried out at the .05 level, and it was found that the data provided no evidence against the null hypothesis. Therefore, individual class estimates of the correlation coefficient were combined to give an overall estimate of $\hat{\rho} = .50$, indicating that about 25 percent of the variance in one measure is predictable by a linear function of the other. The value of 25 percent is neither low enough to make one wash one's hands of one or both measures, nor high enough to give one a supreme feeling of confidence in talking about risk taking as determined by either of these two measures. However, it is clear that there is some common variance, and in that respect, each of the measures remains at least somewhat interesting.

TABLE II

Correlations between the disguised and reported measures of risk taking

	C_1	C_2	C_3	B_1	B_2
n	19	30	9	39	23
$r_{R_zR_{z'}}$	.61	.52	.77	.54	.10

BIBLIOGRAPHY

Terman, L. M. Concept Mastery Test. New York. Psychological Corporation. 1956.

Ziller, R. C. "A Measure of the Gambling Response-set in Objective Tests." Psychometrika. 22: 289-292. September 1957.

PROBLEMS IN THE ASSESSMENT OF GROUP COUNSELING

David G. Zimpfer, University of Rochester

Group counseling is a process for aiding students in their vocational, educational, social, and personal adjustments and development; it uses the medium of small group interaction under the leadership of a trained counselor. This paper will comment critically on the present scene in group counseling.

Group counseling in educational settings was first begun about twenty years ago. It seems to have developed largely out of the desperation of counselors who had large caseloads. When seen as a substitute for individual counseling, two assumptions need to be made, and each is impeding present development. (1) The goals of individual counseling are appropriate also for counseling with groups. This has hampered group counseling developing usefulness and objectives uniquely its own. Typical criteria for individual counseling such as improvement of grades, congruence between real and ideal self concept, and realism of vocational choice, are used also with group counseling. (2) Groups need to be focused on a common problem in order to be able to be treated simultaneously by one counselor. The stress in the literature on group freedom and permissiveness to talk about the real issues which concern members, and on granting members the right to diverge and evolve topics and solutions to problems, has hampered group counseling. The common-problem assumption implies also that the predetermined choice of criteria and outcome measures be applied uniformly to all group members.

This paper focuses on problems in two areas: (1) conceptualization in group counseling, and (2) aspects of research methodology.

Conceptualization

There is no true theory in group counseling to date, only rationales and suggested procedures. Group psychotherapy has theory, and group dynamics possesses research capabilities, but group counseling has not developed a literature of its own. A researcher, in consequence, cannot be sure if the several counselors he may use in an experiment are well qualified in terms of a particular orientation, or whether they will perform similarly.

The issue becomes clear when one examines the positions writers take regarding the group counseling process; should the counselor focus on the individual within the group, or on the dynamics within the total group? The former uses individual counseling style, with the group serving as an incidental resource for support or questioning. The latter involves working with group dynamics to the advantage of the group; the counselor interprets them to the group as an important learning objective of the counseling process. The latter position clearly suggests that the use of a group offers unique advantages not found in other treatments.

The imposition of a single criterion instrument violates the individuality of members. It also hinders progress of the group from one objective to another; these are fundamentals in the counseling process. I would propose that objectives and criteria be thought of as constantly in process, rather than fixed and predetermined. There are no conceptual or research grounds for doing otherwise.

There is an extreme lack of information on what actually went on in the process of counseling in typical research literature. Journal editors should allow space for research contributors to discuss the rationale on which their procedures are based, and to elaborate on the very processes themselves that occurred in their experiments.

The lack of theory and adequate conceptualizing and the withholding of relevant descriptive data in research articles, have resulted in proliferation without concomitant testing of value. Different writers conceive of using one counselor or more than one, groups of three to five members or as many as 30 or more, focus on individuals versus focus on group process, informational versus affective goals. They question what the most appropriate type of leadership is.

Methodology

Whatever the advantages of the group setting in counseling, the outcomes selected for focus are not typically those which draw on these advantages. The grade point criterion is often used instead of criteria which draw mainly on the very properties of groups.

Group counseling outcome research often concludes with null results. Most of the researches have been conducted with relatively short term counseling, eight, ten or sixteen sessions. Perhaps it is unrealistic to expect large positive gains on the criteria often used (gains in self concept or in grade point average) from such relatively short-term counseling experiences. It may not be appropriate to expect eight or sixteen sessions to change atitudes or behaviors built up over years of life experience. The commonly advanced argument that the nature of treatment is correct and that adjusting its duration will suffice to bring about greater change, is a position that has not been tested. There is need for replication studies in which the time variable is the only consciously altered condition.

Given the time limits of most school and college counselors, it may also be unrealistic to expect more than fifteen or so sessions of group counseling. Perhaps more immediate goals, dealing with decisions, plans, and other cognitive outcomes may be more suitable. Even such a limited outcome as a simple commitment to examine the possibility of change may be more reasonable.

Evaluation

Most studies examine results on the criterion measures immediately after counseling. What chance does an individual who has relatively low grades at the outset of several months of counseling have of changing those grades by the posttest? I submit that the process of self-change, teacher-change ("This boy *is* doing better.") and grade-change is not possible by the posttest. In one recent study, (Ofman, 1964), it took three semesters for grades to be affected subsequent to group counseling.[1] Immediate posttest evaluation ignores the possibility of delayed and cumulative effects accruing as a result of counseling.

Too little is known about the process of groups. Many questions exist on group size, transfer of learning, manipulation of group dynamics variables, personality combinations, open versus closed membership, and the relationship of group cohesiveness and morale to the members' right to individuality of outcome. The appropriateness of the instruments commonly used for evaluation, and the proliferation of homemade devices, are also questioned.

My concern is that the potential of group counseling will be squandered by misapplication of procedures and measurement devices, by inadequate conceptualization, by careless designs, or by failing to take advantage of the unique properties of groups.

---

[1] William Ofman. "Evaluation of a Group Counseling Procedure," *Journal of Counseling Psychology*, 1964, 11, 152-159.

ASSESSMENT OF ATTITUDE TOWARD SCHOOL: A PRELIMINARY STUDY

Edward E. Panther, State University College at Buffalo

How well a child does in the first few years of school is determined by intelligence, motivation, physical condition, and environmental circumstances. One important factor is the attitude of the child toward the school situation, but here, the problem in studying attitudes of young children is the lack of adequate assessment devices. Questionnaires, self-rating scales and other methods of assessing adult attitudes are of limited value when used with children.

Purpose

The purpose of this study was to develop an instrument for assessing children's attitudes toward school, and to explore the possibility of using this instrument in assessing the effectiveness of preschool compensatory programs in helping children to develop more positive attitudes toward school.

It was felt that children's attitudes toward school could be assessed through the use of a series of drawings depicting animals in school type or school related situations. The rationale for the use of such an assessment device is similar to the rationale behind projective devices such as the CAT (Children's Apperception Test). The children's responses to the stimulus drawings were assumed to reflect their attitudes toward similar situations. This paper is concerned with the first part of the study, development of the instrument, and a pilot study to attempt to evaluate the responses to the stimulus items.

Method

The stimulus items consist of seven black and white water color and charcoal drawings. Three of the drawings depict animal characters in schoolroom situations: one drawing is of school playroom activity; one is of an outdoor play situation where one character in the drawing is carrying books; one shows the animal characters entering a school building; and one is of a homework situation.

The pilot group consisted of 46 kindergarten children from the Campus School at the State University College at Buffalo. It also included 19 prefirst-grade Negro children from the Buffalo innercity schools, who were enrolled in a summer program for culturally disadvantaged children. Each child was assessed individually by the investigator or an associate. The children were given the following instructions: "I am going to show you some pictures and I would like you to look at them and then make up a story about each of them." The stimulus drawings were then presented one at a time and the responses were written down by the examiner. When responses consisted only of a description of the stimulus drawing, or when the response was vague, the examiner asked questions so that more adequate responses might be elicited. While the form of further questioning was not perfectly consistent, it usually consisted of the questions: "Could you tell me more?", "How do you think the character in the drawing feels?", "How does the story end?", or "What is happening in this picture?".

Each response was rated on a one to five scale. Those responses which indicated that the stimulus situation was seen as very negative, or threatening, were assigned a rating of one. Those which were seen as very positive, pleasant or enjoyable were rated five. A rating of three was given to responses which seemed to be neutral or ambiguous. Ratings of two and four were assigned to responses which were less extreme than one or five responses. The ratings for each of the seven responses for one child were added together so that it was possible for a child to score anywhere between 7 and 35 on the combined rating. For the preliminary study all rating was done by the investigator.

The 46 kindergarten children were in two groups, morning and afternoon classes. The mean scores for these two groups were subjected to the t-test to see if there was a significant difference between the attitude response means. The t-test was also used to compare the means of the Kindergarten groups with the mean of the group of children in the program for the culturally disadvantaged.

Results and Discussion

The results of the preliminary study are presented in Table I.

The results indicate that there was no significant difference between mean ratings of the two Campus School kindergarten groups. The mean rating of the responses of the children in the summer program differed significantly from both Campus School groups. The responses to the stimulus pictures made by children in the summer program for the culturally disadvantaged were rated as more positive than were the responses of the other two groups.

Obviously there is more than one possible interpretation of these results. There seemed to be some differences between the Kindergarten groups and the summer program group in their attitudes toward school as measured by the instrument. Whether or not this is generally true of culturally disadvantaged children as compared with middle class children at this age level remains to be seen. It is also possible that other factors (rater bias, time of year, length of time in a compensatory program, race of examiner) could have affected the results. What is important, is that the results indicate that the assessment device may be useful for assessing groups of children in terms of positive or negative feelings toward school situations.

TABLE I
Means, Standard deviations, and t-ratios for the ratings of the Preliminary Group's Responses

Means and Standard Deviations

Group	Mean	S.D.
A.M. Kindergarten	21.13	3.04
P.M. Kindergarten	21.09	3.46
Summer Program	24.26	2.98

t-ratios between groups

Groups	t-ratios	df
A.M. Kindergarten-P.M. Kindergarten	.040[1]	44
A.M. Kindergarten-Summer Program	3.384[2]	41
P.M. Kindergarten-Summer Program	3.117[2]	39

[1] Not Significant
[2] $p < .01$

Conclusions

The results of this preliminary investigation are encouraging. The assessment device which was developed appears to be useful for investigating children's attitudes toward the school situation. Future research with this instrument, under a more carefully controlled research design, will be undertaken. From experience with the preliminary study, the investigator feels that the usefulness of the instrument with preschool children could be improved by developing a standard set of questions for each card since their responses are frequently limited and additional questioning is often necessary. It will also be necessary to use the instrument with older children. Data will be obtained for first, second and third-grade children, at least. In addition to the findings reported in this study, the investigator has used the instrument successfully as an individual projective device.

Use of the instrument to assess the effectiveness of programs for culturally disadvantaged children requires that pre and post data be obtained for children enrolled in such programs and for a control group of similar children not enrolled in a program. Data from a larger sample of middle-class children must also be obtained.

One question, which has arisen as a result of this preliminary study needs to be carefully investigated. If the findings (that culturally disadvantaged preschool Negro children in compensatory programs have more positive attitudes toward school than middle-class children) are verified by future research, a longitudinal follow-up study will need to be undertaken. This would need to try to discover how and when their attitude toward school may change as a result of experiences in the first three or four years of school.

CHARACTERISTICS OF LOW AND HIGH PERFORMING ELEMENTARY PUPILS

John H. Rosenbach & Robert F. McMorris, State University of New York at Albany; & William Smith, Guilderland Central Schools

Two years ago we began a longitudinal study directed toward a redefinition of what have typically been called *under-* and *overachievement.* The study arose from a concern that the theoretical aspects upon which these concepts are based is inadequate. For example, Thorndike has cogently argued that underachievement and its counterpart overachievement can be reduced to prediction errors; underachievement represents a matter of overprediction while overachievement results from underprediction.

The present study represents an effort to reconceptualize the concepts of under- and overachievement by starting with the assumption that learning (or achievement) is an hypothetical construct. As such, the amount, rate, and quantity of an individual's learning are inferred from his behavior. The inferences made will depend upon the kinds of behavior samples that are chosen for observation. To illustrate this position, a pupil's achievement test represents but one index from which achievement or learning may be inferred. Teacher evaluation, as based on grades, for example, represents a different index. In each instance the index of achievement will depend on a number of variables, such as motivation, the behavior sample, and the differential weights given to various factors. It is to be expected, therefore, that the two measures will not coincide.

The present study was designed to identify children at the elementary school level who showed a marked discrepancy between learning as measured by standardized achievement tests and school performance as perceived by teachers and reflected in teacher grades. In addition various cognitive and noncognitive measures were administered in an attempt to identify personal characteristics that differentiated low performing pupils from high performing pupils.

In order to select discrepant groups (low performers and high performers) a two-stage regression approach, as suggested by Farquhar and Payne, was used. Fourth and fifth grade ITBS scores were used as predictors and teacher grades as the predicted variable.

The selection model was applied separately for the sexes. Initial analysis of the data was done on the basis of mean differences between several variables for the two pairs of discrepant groups and was reported in detail elsewhere. Briefly, general support was found for the hypotheses that:

1. High performing pupils will show greater need for school achievement than low performing pupils
2. High performing students will show more positive orientation toward school tasks than will low performing students.

The data which we are reporting on now, relate first to the reliability of the constructs under consideration, high and low performance;—two which have methodological implications for studies of this kind. Secondly, the data relate to the identification of variables which in turn relate to this learning-performance discrepancy, and the correlations that were found among the several variables.

In summary, our findings suggest that more refined methods to identify high and low performers are necessary in order to produce a more reliable construct. Unfortunately, in spite of a great number of studies dealing with underachievement, it appears that little effort has been exercised in the past to study the reliability of under- and overachievement. If the numerous, and many times, arbitrary operational definitions used in the selection of subjects in the study of the phenomenon cannot be demonstrated to yield reasonably reliable results, it would seem that the value of these findings must be seriously questioned.

Second, our findings give support to Bloom's contention of the stability of achievement over time, and furthermore suggest that a good deal of information collected in the schools is highly redundant.

And thirdly, the reasonably high correlations of such noncognitive measures as the Personal Values Inventory and the Work Habits Rating Scale, with cognitive variables, suggest that these measures are of value for further study.

THE MAJOR PURPOSES AND FUNCTIONS OF SUPERVISION AS PERCEIVED BY NEW JERSEY PUBLIC SCHOOL SUPERINTENDENTS, SUPERVISORS, AND BUILDING PRINCIPALS

Geraldine Frances Cleminson, Fordham University

This study investigated perceptions of supervision expressed by all New Jersey public school superintendents, supervisors, secondary principals, and elementary principals when assessed according to the democratic, authoritarian, and laissez-faire philosophies of supervision and five independent variables: sex, total years of professional service, predominance of experience in administration or supervision, predominance of experience at the elementary or secondary school level, and possession or lack of the doctorate.

The instrument developed for this study was patterned according to the Likert attitude scale techniques; 90 randomly arranged items, based on a comprehensive review of the literature, were experimentally defined as democratic, authoritarian, or laissez-faire by a national and state jury of experts. Items which achieved more than 80 percent juror agreement were administered a proportional, representative random sample of 250, or 10 percent, of the total State sample of 2,514 superintendents, supervisors, secondary principals, and elementary principals in order to determine item discrimination. In responding to the preliminary instrument of 67 items, subjects could "strongly agree," "agree," be "uncertain," "disagree," or "strongly disagree." Eighty percent of the responses were usable. All items except one revealed a *t* score of .001. Those 12 items on the democratic scale, the authoritarian scale, and the laissez-faire scale, representing the highest *t* scores constituted the 36-item final instrument.

The final instrument was administered the total State sample, exclusive of participants in the refinement of the instrument. Final instrument respondents had the same options in item response as had the preliminary sample. Of the 1,881 (83 percent) of the final responses received, 1,597 (85 percent) were usable. Discrepant response distributions, necessitated adjustments in the statistical treatment of the data.

Analysis of variance techniques was applied to the total scale scores for each of the four professional role categories. Differences in responses among superintendents, supervisors, secondary principals, and elementary principals were considered significant at the .05 level. Chi-square techniques were employed in analyses of the responses to the individual scale items; relationships between professional role and responses, likewise, were considered significant at the .05 level.

The null hypotheses tested revealed the following:

1. Females were more democratic than males, and supervisors more democratic than administrators.
2. Female administrators and male supervisors, in their strong endorsement of democratic supervision, accounted for the significant interaction of sex and professional role.
3. Male supervisors and female administrators manifested greater approbation of authoritarian supervision than female supervisors and male administrators.
4. Female respondents denoted stronger commitment to laissez-faire supervision.
5. As the total years of professional service increased, commitment to either authoritarian or laissez-faire supervision increased. Total years of professional service was not significant in respondents' commitment to democratic supervision.
6. Predominance of experience in administration or supervision was significant only in the respondents' perception of democratic supervision. Supervisors with predominant administrative experience and administrators with predominant supervisory experience indicated the strongest endorsement of democratic supervision; administrators with administrative background denoted least support of the democratic view.
7. Predominance of experience at the elementary or secondary level was not significant in perception of democratic, authoritarian, or laissez-faire supervision.
8. Respondents possessing the doctorate were more democratic; respondents lacking the doctorate tended to support authoritarian supervision.
9. Significant relationships existed between the professional role and 5 of the 12 democratic, 4 of the 12 authoritarian, and 5 of the 12 laissez-faire items.

A STATISTICAL MODEL OF CONFORMITY-DEVIANCE

Joan Roos Egner, Cornell University

Central to the problem of the study is the relationship between future administrators and their organization and reference groups. Graduate students training for administrative positions have several groups to which they belong and/or may use as a referent. One such group is the Inter-University Project II, The Administrative Internship, developed cooperatively by the Universities of Buffalo, Cornell, Rochester, and Syracuse. As a demonstration program for preparation of educational leaders, the Administrative Internship has utilized varied procedures and concepts to assist in the development of administrators of quality. The relationship of the members of an internship project, educational organizations, reference groups, and groups of the larger culture to the conformity-deviance variable, appears to have significance as one seeks to appraise a program for administrative training and to evaluate that program's impact on educational administration.

The research seeks to investigate the service provided by an internship program to administrative preparation. By ascertaining the intern's degree of conformity to the social norms of the society he will serve and lead. His conformity or deviance from university, school, or societal norms may be an indication of how well he has learned general cultural norms and the extent to which he has internalized norms particular to his profession.

Purpose of the study was to establish and analyze statistical norm patterns for interns in educational administration. In addition, statistical norms were established for sponsoring university faculty and administrators, nonsponsoring administrators, and graduate students in educational administration. These norms were then compared to internship project norms. The investigation was concerned with a statistical model of conformity-deviance. That is, analysis centered around the frequency of various types of responses rather than the content of the responses.

The total population ($N = 203$) provided a source of study for relationships between conformity and age, sensitivity training, administrative experience, hierarchical level, and organization size. Data were compared with a representative sample of the United States population and selected subgroups.

Data were gathered by use of the Tomkins-Horn Picture Arrangement Test (PAT) and a data sheet. The PAT and data sheet were mailed to, and self administered by former interns (1962-65), sponsoring and nonsponsoring administrators, university faculty, and graduate students in educational administration. The PAT and data sheet were mailed to the 1965-66 interns before they organized as a group for summer session sensitivity training. After sensitivity experience, the interns were administered the test in a group setting. Six months later, during the interns' field experience, the interns again completed the PAT.

The tests were scored according to standardized instructions. Statistical norms were determined for intern groups within project years, universities and the total project, university faculty, sponsoring and nonsponsoring administrators, and graduate students of educational administration.

At this time, analysis of all data is not complete. The main focus thus far has been on analysis of group social norms using the statistical norm concept developed by Miner. That is, if 50 percent or more of the group members select a certain response arrangement on the PAT it is considered a norm for that group. The maximum number of norms that can be scored is 25. Miner's representative sample of the United States population yielded a statistical norm of 9.

The four intern groups from 1962 through 1965 showed strong normative patterns. The 1962 groups had a statistical norm of 18 as did the 1963 group. Interns of 1964 had a norm of 20 with the 1965 group a mean of 21. Homogeniety of the intern group was further indicated by noting an identical response arrangement of 13 was maintained over the four years by the four groups. When these norms of each year's educational administration internship group and their total norm response are compared to the United States norm of 9, as well as a comparison with a sample of business executives with a norm of 15, it appears that a strong normative pattern is defined within the administrative internship.

The research sample is homogeneous on the conformity variable, as defined, and is more conforming than the United States representative sample. Homogeniety of the group and high conformity to societal norms merits further analysis. There may be a strong tendency for educational organizations to select and retain their members, or for members to select these groups using the implicit or explicit behavior patterns as a guide. Further, each group was composed of members who had graduate level education. The potential for conformity may be higher in these groups since intellectual capabilities would facilitate internalization of standards and thus a greater number of conforming societal responses could be acquired. Additional findings will be gleaned when tests of significance are applied to data. Variables of age, sex, educational and experiential level will be organized in a matrix with PAT data to determine relationships.

At this point in analysis, findings indicate the intern population as a group conforms to social norms of the population they serve and lead. They reflect the norms of their culture as well as their profession.

"PERCEIVED NEED DEFICIENCIES OF TEACHERS AND ADMINISTRATORS: A PROPOSAL FOR RESTRUCTURING TEACHER ROLES"

Francis M. Trusty & Thomas J. Sergiovanni,
University of Rochester

This study investigated need fulfillment deficiencies for teachers and administrators in a school system (n=233). Maslow's theory of human motivation served as the conceptual framework. Lyman Porter's industrial management study provided the point of departure. This study utilized a variation of Porter's questionnaire and analyzed the role of five basic needs: security, social, esteem, autonomy, and self-acutalization. The general hypotheses of this investigation were: (1) there is no difference in the perceived needs of educators in one age group as compared with educators in other age groups, (2) there is no difference in the perceived needs of educators in one experience group as compared with educators in other experience groups, (3) there is no difference in the perceived needs of male educators as compared with female educators, and (4) there is no difference in the perceived needs of educators in one professional role as compared with educators in other professional roles.

In general, the need deficiencies of educators in this study varied according to age, sex, and professional role. Accordingly, educators value differently opportunities for achieving economic security, developing social relations, increasing one's self-esteem, having autonomy or becoming self-acutalized. Specifically, the overall response by age groups (H_1) indicates that the 20-24 age group has the least and the 25-35 age group has the greatest need deficiencies. In general, the hypothesis (H_2) that years of experience *would not* make a difference was confirmed. The hypothesis related to sex of respondents (H_3) was rejected. Women appear to have a greater need deficiency in regard to security than men. Further, women appear more satisfied with their professional role than men. Professional role (H_4) appears to be the variable most significantly associated with need deficiencies of educators. Assuming that need deficiencies are an index of dissatisfaction, it seems that the role of the junior-senior high school teacher affords less opportunity for satisfaction than other professional roles. It appears that the job of administrator permits the greatest opportunities for experiencing need fulfillment at all levels. Elementary teachers compare favorably with administrators.

This study suggests that increased opportunities for professional advancement of teachers is needed if they are to experience more need fulfillment at higher levels of the Maslow hierarchy. One approach to increasing professional opportunities is to individualize the process by which teachers' needs are satisfied through the creation of a separate hierarchy for teachers. A proposal for restructuring teacher roles is made based on the premise that as teachers increase in professional competence and maturity they should, and need to, be given more responsibility for their professional behavior and opportunity for personal and professional growth. Four levels of teaching are proposed, each reflecting increasing competence on the part of the teacher. A new teacher would function as an *Intern* (1.00 salary index) with a small teaching load and cooperatively work with teachers at the Fellow and Associate level. The *Fellow* (1.75 salary index) would become increasingly involved in guiding the learning of students and participating in the allocation of educational resources. Fellows would constitute the nucleus of the professional teaching core. The *Associate* (2.75 index) would share the responsibility for the professional growth of Interns and Fellows. In addition, the Associate would have maximum exposure to students in individual and cooperative teaching relationships. The Associate would also work cooperatively with parents or teachers on educational problems. Movement from the associate level can take place in two directions: Scholar or Colleague. These positions would be considered equal in rank and pay (index 4.00) but would provide a different emphasis in terms of responsibility. The *Scholar* would serve as an educational clinician, involved in curriculum research and development, teaching jointly with Fellows and Interns, and working with the exceptional

student. The *Colleague* would possess strength in human and organizational skills. The Colleague also would have a strong voice in curriculum development, policy formation, and would maintain close ties with the school administration, the school board, and the local community.

PERSONNEL PRACTICES IN SPECIAL EDUCATION

Robert E. Mulligan, Public Schools, Fords, New Jersey

The American Republic, from its very beginning, has indicated over and over again its commitment to universal education. This includes the education of mentally retarded trainable children. Notwithstanding the need, education for the severely mentally retarded, until recently, was almost an unknown quantity in public schools. And the need has been great. Masland contended that "it is unlikely that in this country there is any other form of disability which equals impairment of mental efficiency. One can say that, with the possible exception of mental illness, mental retardation is the most significant handicap of our present society."[1] Among the mentally deficient we encounter in schools the trainable child is most handicapped.

How many of these youngsters are there? According to a national survey made by Kirk[2] in 1957 the incidence of trainable children was between one and two youngsters for every thousand children of school age. Surveys conducted by the Department of Public Instruction in Michigan[3] and Illinois[4] revealed a similar frequency.

Perhaps the low incidence in terms of percentages explains why, until recently, public education demonstrated what Kanner[5] described as a mixture of taunting neglect and good-natured toleration. A dismal commentary on the manner in which we have implemented our commitments to trainable children is in the indictment of educators by Lloyd Dunn[6] who believed that the recent national trend to educate trainables "resulted mainly from the demands of parents . . . rather than from the desire of educators to assume responsibility for this group." More stinging is Magnifico's[7] accusation that there are many among us who do not admit it publicly, but who believe that the cost of training the severely mentally deficient is a waste of public resources.

On the positive side, an enormous effort has been made since 1950 to provide trainable children with a suitable educational program. However, this like so many other ventures in education, was begun rich in intention and poverty stricken in know-how. Research studies in teaching methods, attainable behavioral goals, suitable materials, and appropriate facilities for the trainable child are sparse indeed. Research studies which we might describe as administrative in nature have been confined to headcounting and certification criteria. On the tenth anniversary of mandatory legislation on behalf of trainables in New Jersey a study was made to investigate teacher-personnel practices used by administrators in developing programs for trainables in the public schools of that State.

The personnel practices with which this study was concerned were in three general areas: practices used in obtaining teachers of trainables, practices used in developing those teachers, and appraisal practices. The study attempted to determine the extent of use of the practices, the value attributed to the practices by those who used them and the reasons for the nonuse of the practices.

A check list was constructed from an analysis of the professional literature related to the principles of personnel administration and professional empirical sources. Eighty-six practices were included in the check list.

Each superintendent was asked, in responding to the check list, to indicate whether or not he considered the practices he used effective. For those personnel practices he did not use, he was asked to check one of six reasons why he did not use the practice. The suggested reasons for not using the practices were: 1) no worth, 2) no need, 3) staff problem, 4) no authority, 5) nonexistence, and 6) no consideration.

---

[1] Richard L. Masland, "The Prevention of Mental Retardation, A Survey of Research," *American Journal of Mental Deficiency,* LXII (May, 1958), p. 991.

[2] S. A. Kirk, *Public School Provisions for Severely Retarded Children. A Survey of Practices in the United States,* Special Report to the New York State Inter-departmental Health Resources Board (July, 1957).

[3] R. J. Guenther, Final Report: *A Michigan Demonstration Research Project for the Severely Mentally Retarded* (Lansing: Michigan Department of Public Instruction, August, 1956).

[4] H. Goldstein, *Report Number II of Study Projects for Trainable Mentally Handicapped Children* (Springfield: Illinois Department of Public Instruction, January, 1956).

[5] Leo Kanner, "Itard, Seguin, Howe, Three Pioneers in the Education of Retarded Children," *American Journal of Mental Deficiency,* LXV (July, 1960), p. 2.

[6] Lloyd E. Dunn (ed.), *Exceptional Children in the Schools* (New York: Holt, Rinehart and Winston, Inc., 1963), p. 129.

[7] Leonard X. Magnifico, *Education for the Exceptional Child* (New York: Longman's Green and Company, 1958), p. 3.

The population of this study consisted of the 140 superintendents of schools whose districts conducted approved classes for trainables during the 1964-1965 school year. Three kinds of districts were included, 1) elementary school districts, 2) K-12 districts and 3) regional secondary school districts. One hundred thirty-seven (97.86%) check lists were returned and 135 (96.43%) were considered usable.

In considering the use and nonuse of the selected practices, the value attributed to them, and the reasons for their nonuse, the effect of the following variables was determined: 1) the total expenditure per weighted pupil per district, and 2) the number of professional staff members for 1,000 weighted pupils. Both variables used in the study involved a weighting process employed by the New Jersey Education Association to facilitate a comparison of school districts which were structurally different. Since the districts in the study were structurally of three kinds, the weighting process provided reasonably comparable data when an attempt was made to compare the effort made by each district. The weighting process has been widely accepted as valid for such comparative purposes. It has been used by the New Jersey Tax Policy Commission of 1963, the New Jersey Department of Education, the New Jersey Federated Boards of Education as well as many local school districts.

The responses to the check list were analyzed in terms of frequencies and percentages. These data were subsequently subjected to a t-test of the biserial correlation coefficient in the cases of responses to use and evaluation of the selected practices. The chi-square test was used with data related to the reasons for not using the practices. The study revealed that of the 86 practices 58 (67.44%) were used by a majority of the respondents.

This study was initially predicated on a series of null hypotheses which in essence assumed that the use of effective personnel practices was not related to the amount of money which the district spent and not related to the staff to pupil ratio in each district. Generally, the conclusions we must infer from the findings are that the null hypotheses of no relationship must be accepted. Of the 86 practices the use of only 16 percent was significantly related to the variable of expenditure and the use of only 14 percent was significantly related to the variable of staff to pupil ratio.

Neither of the two variables employed in the study seemed to be appreciably related to the value attributed to the practices by those who used them. Of the 86 practices the evaluation of only 4 were significantly related to expenditure and the evaluation of only 3 of the 86 were related significantly to the staff to pupil ratio.

Because of the highly skewed distribution of responses to the reasons for the nonuse of the 86 practices, the calculation of chi-square could only be made in 40 of the 86 situations. Of these 40 practices the reasons for the nonuse of 29 percent were related to staff to pupil ratio, but only 5 percent were related significantly to the expenditure variable.

Analyses of the responses of the superintendents to the check list suggested two general conclusions: 1) the amount of money a district spends and the staff to pupil ratio related less significantly to the use and reasons for nonuse of personnel practices than was originally assumed, 2) the large number of respondents who gave as a reason for nonuse "no consideration has been given to the use of this practice" would indicate a possible lack of familiarity with the literature in personnel administration.

INSTRUCTIONAL MODE, PERSONALITY CHARACTERISTICS, AND ACHIEVEMENT: OVERVIEW OF THE RESEARCH AND DEVELOPMENT OF INSTRUCTIONAL MATERIALS

Richard E. Ripple, Cornell University

The Learning Structures Project (LSP I) came into being on July 1, 1964, financed by a grant from the Cooperative Research Branch of the United States Office of Education. The Project is scheduled for completion on July 1, 1967. Although the final report has not been written, all of our data are analyzed. Several interim reports have been published and papers have been read at conferences, but this session represents our first effort to give a complete (although abbreviated) report of our results.

In order to give some idea of the dimensions of the project, approximately 5,000 eighth-grade students from 22 junior high schools in New York, New Jersey, and Pennsylvania participated in the pilot year, 1964-65. This year was spent in refining our data gathering instruments, instructional materials, and other research procedures. The experimental year, 1965-66, involved the participation of approximately 1,200 eighth-grade students from 20 of the original 22 junior high schools in the same three states.

Rationale for Selection of Learner Characteristics and Specific Research Objectives

The general problem to which our research addressed itself can be summed up with this question: "What relationship, if any, exists between selected personality characteristics of students and the relative degree of success they have in learning from programed instructional materials and from conventionally structured learning tasks?"

The structure of the task involved in learning from programed material differs from the conventionally structured learning task in several respects. In learning from programed materials:

1. the learner works by himself at his own rate of speed;
2. the learner receives immediate feedback about the correctness or incorrectness of his responses;
3. the subject matter content of the lesson is arranged in a series of very small steps which tend to repeat the essentials of the lesson; not only is the content repeated for the learner, but the entire manner of presentation is repetitious.
4. the learner's freedom of response is restricted; that is, the response which the program will allow from the learner is extremely limited, and the learner's thinking is forced to conform severely to the structure of the program.

Our interest in anxiety stemmed from its well recognized effect of interfering with learning in the classroom. Anxiety is defined as a stable response tendency, composed of physiological reactions, and is realized by the individual as a generalized state of unpleasantness. The anxious child is characterized as excessively dependent upon external support in the learning situation and is easily disconcerted by evaluative and other adverse comments made by his teacher. We reasoned that the structure of the task involved in learning from programed material would be less likely to arouse the debilitating anxiety responses in students than would the more conventionally structured learning task. Consequently, we expected that learners scoring high on tests of anxiety would do relatively better in the programed learning task situation.

We conceived of exhibitionism as the degree of an individual's positive attitude toward showing himself and his products to an audience—a motive wherein the individual's goal is approval for public acts. The line of reasoning is clear. The structure of the task involved in learning from programed material does not permit an individual to exhibit his products to an audience as much as does the more conventionally structured learning task. Hence, we expected that learners scoring low on tests of exhibitionism would do relatively better in programed learning.

Compulsivity has been defined as the tendency to perform some action, even when it is known to be unnecessary and absurd. The compulsive individual is described as one who demands that situations have a high degree of order. Again our line of reasoning, although naive in retrospect, is clear. The structure of the task involved in learning from programed instruction is characterized by greater repetition and order than is the conventionally structured learning task. We expected that learners scoring high on tests of compulsiveness would do relatively better in the programed learning task situation.

Finally, we took Guilford's distinction between convergent and divergent thinking as a point of departure in identifying the relevance of creativity. Convergent thinking leads to the "right answer" which can be determined from the information given. Divergent thinking is defined as the kind which goes off in different directions. It makes possible changes of direction in problem solving and leads to a diversity of answers where more than one answer is appropriate. We reasoned that the structure of the task involved in learning from programed material would inhibit (and therefore interfere with learning) of the individual who is characterized in his thinking style by divergent rather than convergent thinking. We expected that learners scoring high on tests of convergent as compared to divergent thinking would do relatively better in the programed learning task situation.

Procedures

Subjects, Treatment Groups, and Method

In each of 20 junior high schools, two eighth-grade English classes roughly equivalent in intelligence and sex distribution were identified. One class was randomly assigned to a programed instructional condition, the other to a conventional instruction condition. Similar instruction in vocabulary development by means of affix and root analysis and use of context were presented by the two instructional conditions in ten regularly scheduled periods. Each instructional treatment consisted of approximately 600 subjects. The specific nature of the instructional treatments will be described later.

Analysis

For each analysis, subjects in *each school* were placed into one of 16 subgroups based on dichotomization of the independent variables: sex, intelligence (high, low), level on the personality variable under consideration (high, low), and treatment condition (programed, conventional). Then the classifications were collapsed for subjects across schools. Scores on the four vocabulary criterion tests served as dependent variables. Four-way factorial analysis of variance were carried out for all personality variable and criterion test combinations. The main effects of sex, intelligence, treatment condition, and each of the several personality variables together with all interactions were obtained and tested for significance. The tests of significance for the interaction of each personality variable and treatment condition on each of the four dependent variables constituted the appropriate hypothesis tests addressing the major research questions raised by this study.

Development of Instructional Materials

Programed Instructional Treatment Condition

The original programed instructional materials used in this research were devised by Glock and Schepman. The content of the program can be described as vocabulary development by means of affix and root analysis and use of context. More specifically the program taught word elements, affixes, and combining forms which have fixed and invariant values.

The original materials went through the following development. First, the program was tried out with 3 average students and 3 above average students individually at the eighth grade level. Based on the responses of these students, analysis of the program by the research staff, and analysis by a special consultant, modifications were made. These modifications were intended to increase the interest of the materials for eighth-grade students, manipulate the error rate on some frames, and redesign the form of the program (including the design and manufacture of a program holder). In January and March of 1965 the program was administered to students in ten schools with two or more classes in each school. Approximately 3,000 students were involved in these trials. After each trial, modifications in the program were made. Error rates per item and per lesson were obtained. In no case did the error rate per item exceed 15 percent. The average error rate per lesson was approximately 5 percent.

The final version of the Improving Reading Vocabulary or IRV Program was in the form of 10 lessons —each lesson in booklet form printed on 5 x 8 paper. The program was in a linear format of the constructed response type. Although the program was aimed only at improving reading vocabulary, a pronunciation guide was incorporated in the form of a phonetic, hyphenated version of every new word placed in parentheses immediately after its first appearance in the program.

DEVELOPMENT OF THE CRITERION TEST LEARNING STRUCTURES PROJECT

Jason Millman & M. D. Glock, Cornell University

There were three stages in the development of the criterion test. In Stage I, 189 items were written to measure the objectives of the instructional unit. The purpose of Stage II was to reduce the number of items in the pool by eliminating those that were very easy or very difficult. Stage III culminated in the production of the final criterion test.

Stage I: Writing the Original Items

The vocabulary instruction was based on the following objectives:

1. To define 23 word elements.
2. To define 118 words.
3. To analyze and break down a word into two elements.
4. To recognize the word element which was defined earlier in the lesson.
5. To determine clues pertinent to the remaining word element(s).
6. To ascertain the salience of these clues for the meaning of the remaining element(s).
7. To generate hypothesis concerning interrelationships between the salient clues for the unknown element(s) and the knowledge that the individual has of the known element.
8. To select the hypothesis (a definition) of the unknown word element on the basis of the clues and the knowledge of the known word element.
9. To organize the results of this process into a logical written unit for the answer sheet.
10. To ascertain salient contextual clues.

Items for the criterion test were constructed to measure these objectives as adequately as possible within the limits of school time allowed us. The 20 item types have been grouped into four major categories for ease of discussion. It should be noted, however, that these four a priori categories were not substantiated by a factor analysis.

Stage II: Reducing the Size of the Item Pool

Each of the 189 items representing the 20 item types was placed in one of seven "forms." These "forms" were administered in ten schools to pupils who had completed the instructional unit using the programed materials. Schools were chosen from among those used the following year in the experiment proper. Each pupil was administered two "forms." An average of 371 pupils (range 300-398) answered each item.

This test administration had two purposes. First, it was necessary to obtain preliminary data concerning the amount of time needed to complete items of a certain type. Secondly, it provided an opportunity to determine whether the directions for the various item types were satisfactory. Original directions were found to be adequate and there was need for little revision.

Pupils appeared to make intelligent responses to items of 19 types. Representative items of each of these types were retained in the reduced pool which I shall describe shortly. However, the nine items in Type 17 were eliminated without consideration of item statistic data. It appeared that this type of item was not measuring an objective relevant to the instructional materials. In addition to these nine items, 77 more were eliminated from the original 189 item pool, leaving 103 items in the reduced item pool. A number of the 77 items were eliminated because the item difficulty index seemed unusually high or low. No fixed cut-off values were used.

Other items were rejected because a majority of pupils either knew or did not know, as the case required, certain key words in the item related to a correct answer. For example, in Type 10 it was assumed that pupils knew the meaning of some untaught words and/or did not know (prior to taking the question) the meaning of other untaught words. This assumption was treated as a hypothesis and tested by administering to a randomly selected few in each classroom a list of these words (in lieu of taking the two "forms") and by asking them to supply definitions. The third basis for item elimination was the adequacy of sampling for an item type. If the number of items for a type was excessive when compared to the other types, one or more were deleted.

Discrimination indexes were determined by subtracting the proportion of pupils passing the item in the bottom half from the proportion of pupils passing the item in the top half. The total score on the "form" was used as an interval criterion. This discrimination index was not referred to in deciding which items to eliminate.

Although the median discrimination index is slightly lower for the rejected group of items (.30 vs. .37), this resulted from the greater number of very hard and very easy items in the rejected group rather than to any deliberate selection on the discrimination characteristic.

Stage III: The Final Criterion Instruments

The 103 items in the reduced pool were administered to pupils in four schools different from those used in Stage II. These schools were used in the experiment proper during the next year.

Five new "forms" were developed from the 103 items and each pupil received two. Since all possible pairs of "forms" were administered, and since all the items of one item type were contained in the same "form", it was possible to compute correlation coefficients betwen total scores on each of the 19 item types. A pupil's score for an item type was equal to the number of the items of that type which he answered correctly. Each correlation was based on a number of pupils ranging from 79 to 98.

The first eight item types involve practically pure recall of learned material; the first three of word elements and the other five of taught words. The next seven item types require the pupil to recall meanings of taught word elements or words, and use these meanings with untaught materials (contextual clues, definitions, etc.) provided within the context of the item to determine the meanings of either new words (Types 4, 9, 11, and 13) or taught words (14, 15, and 19). Successful performance on item types 10, 12, 18, and 20 require the ability to obtain meanings of new words from various clues in the item itself. A list of all meanings of the word elements studied would be of no help in solving these items. Item 20 differs from the other three in that to receive a high score a pupil must generate many possible hypotheses about the meaning of a new, untaught word. A verbal fluency, divergent thinking ability, would no doubt be of help here.

The above analysis suggests that the first factor is a recall (or recognition) of taught material. The pupil who diligently memorized all word elements and their meanings would do well on this factor. Factor two appears to involve a transfer skill; particularly to learn the meanings of new words given a context or definition of words or elements in common with the new word. Factor three is similar to Factor two in this regard with the addition that the ability to generate many hypotheses is more important.

Item types one and eleven, a multiple-choice variety, were given preference over item types seven, eight, and sixteen because the latter three types were open ended items like three and six. These latter two types had already been selected as partial measures of the first factor. Because much mental effort and emotional attachment were associated with the second groupings of seven item types, we were reluctant to exclude all of these items from the final criterion test. The four item types selected for the recall subtest seem a reasonable compromise between optimum measurement of Factor A and a cross section of recall–recognition type tasks and item forms.

Although the Generation of Hypotheses subtest is composed of only six items, a total score on this subtest could range from zero to 12. That is because a pupil was given one-half a point for each *different* correct answer supplied for each of the six questions. No more than four answers, however, were counted for any one question.

The total score on the criterion test was a simple sum of the scores on the three subtests. The maximum possible score on the total was 37 (that is 12 + 13 + 12).

Thirty-five minutes were permitted for actual working time on the final criterion test. This proved ample for virtually every pupil. It would be fair to say the criterion test was unspeeded.

The eight item types used in the final criterion test contained 43 out of the 103 items in the reduced item pool. Of these 43 items, 31 were retained. The 12 items (43 — 31) rejected were chosen primarily to avoid duplication of the word elements being questioned and to provide a representative sampling of content from the ten lessons. Two of these items were eliminated because the extraneous content in which the items were embedded was highly similar to content found in other items being retained.

INSTRUCTIONAL MODE, ANXIETY, EXHIBITIONISM, AND ACHIEVEMENT

Robert P. O'Reilly, Cornell University

This section of the report of the *Learning Structures Project*[1] presents the results of the evaluation of the hypotheses relating anxiety and exhibitionism to differential success in learning from programed and conventional modes of instruction.

Construction of the Instruments

The objective of the development of the School Anxiety Scale (SAS) was to construct a number of specific subtests which would be more pertinent to the measurement of anxiety in the school situation, than existing anxiety scales.[2] The procedure for development involved three major stages. In the first stage, original items and items selected from other anxiety inventories were grouped into logical stimulus dimensions. The majority of these stimulus dimensions referred to classroom or school related situations judged to have major threat potential for the student. The initial form of the SAS contained 324 items, classified into 31 dimensions (subtests), and presented in four different test formats.

In the second stage, the initial form of the SAS was administered to 300 eighth graders. On the basis of this administration, 23 of the proposed 31 dimensions were factor analyzed. Both a principal components analysis and a varimax rotation were accomplished for each proposed subtest. Subtests were then selected for a revised and shorter version of the SAS on the basis of the amount of item variance accounted for by a given factor in the rotated factor matrix, and the logical consistency of the items within the rotated factor.[3] Items within a given subtest were rejected or retained on the basis of their factor loadings.

The procedures used in stage two were then repeated in stage three with another sample of 300 eighth graders, and the SAS was then reduced to its present form and size. The final form of the SAS is composed of seven subtests and a total of 60 items, phrased in the form of questions ("true-false"). An index of the degree of anxiety for each subtest is obtained by summing the number of appropriate responses. The total anxiety score is the unweighted sum of appropriate responses to each of the subtests.[4]

The procedures used in the development of the Exhibitionism Scale (ES) were essentially the same as those used for the SAS. However, the results of the varimax rotations at stages two and three were disappointing in that logically and statistically homogeneous groupings of items (subtests) were not obtained. The recourse to this was to compose the final test of those items which had the highest loadings in the principal components analyses at stage three.

The final form of the ES is composed of a total of 45 items distributed over three different test formats. Scores from the administration of the ES during the final (experimental) year of the research had a mean of 24.81, and a standard deviation of 7.55 (N = 1136). The mean discrimination index of the 45 items was 27.13. Internal reliability of ES scores, as estimated by the Kuder-Richardson Formula 20, was .85.

Findings From the Analyses for Subjects Classified on Anxiety and Exhibitionism

Four-way analyses of variance were calculated for *Ss* classified on each of the anxiety subtest scores, total anxiety score, and the other independent variables. Thus a total of 32 analyses were required for anxiety, one for each anxiety score on each of the four achievement test scores. Since there is only a single index of exhibitionism, only four analyses were required in this case. Means and tests of significance for the main effects of each of the independent variables are reported. Of all the double interactions calculated, only the means and tests of significance of the interactions of personality and instructional condition are reported.[5]

---

[1] The Learning Structures Project is a three-year research project, sponsored by the U.S. Office of Education, and directed by Professors Richard Ripple, Marvin Glock, and Jason Millman, School of Education, Cornell University.

[2] Cf. Ruebush, B. K. Anxiety. In H. Stevenson (Ed.), *Child Psychology, the Sixty-Second Yearbook of the NSSE.* Chicago: University of Chicago Press, 1963, pp. 465-516.

[3] Practical limitations during the final year of the research required that the test take no longer than 20 minutes to complete.

[4] Two items are each scored for two separate subtests. Thus the total possible anxiety score is 62.

Main Effects of the Independent Variables in the Analyses for Anxiety

Comparisons of the criterion test means for sex indicate that the differences in the means for boys and girls on the four criterion scores are very slight. None of these comparisons reach significance. The 32 comparisons for mental age show that the criterion test means for high mental age are consistently and significantly higher than those shown for low mental age. For instructional condition, the means for recall, hypotheses making and total criterion are significantly higher in the conventional condition than in the programed condition, in all analyses. The means for transfer are also higher in the conventional condition than in the programed condition, but are not significantly different. The criterion tests means for each of the anxiety subtests and total anxiety score indicate that high anxiety is associated with lower mean scores on the criterion test in nearly all analyses. Two exceptions to this are the comparisons for *peer anxiety* on recall, and *teacher anxiety* on recall. However, the difference between each of these pairs of means is very minute. Significantly higher mean criterion test scores were obtained for low anxious *Ss* in the following analyses: (1) *generalized school anxiety* on transfer; (2) *general classroom anxiety* on recall, transfer, and total criterion; (3) *parental pressure for achievement* on hypotheses making and total criterion; (4) *test anxiety* on transfer; and (5) *total anxiety* score on transfer and recall.

Interactions of Anxiety and Instructional Condition

None of the interactions of anxiety and instructional condition on the four criterion test scores are significant in each of the analyses utilizing a different anxiety score. For all comparisons, the lines connecting the criterion test means for subjects grouped high and low on a particular anxiety score, within each instructional condition, are parallel or nearly parallel. Also present is the previously noted negative effect of anxiety on the criterion test means, and the positive effect of the conventional condition on the criterion test means. In summarizing these results, it is evident that there is no tendency (significant or otherwise) for the data to support the hypotheses concerning the interaction of anxiety and instructional condition on achievement.

Main Effects of the Independent Variables in the Analyses for Exhibitionism

The results of the analyses of variance for the main effects of sex, mental age, and instructional condition on the criterion test scores are consistent with the results obtained for these independent variables in the analyses for anxiety. The effect for sex on all four criterion test scores is nonsignificant, with girls obtaining slightly higher mean scores than boys for recall, hypotheses making, and total criterion. The main effect for mental age is significant for all criterion test scores, with the higher means shown for high mental age in each case. The means for instructional condition indicate that on the average, *Ss* in the conventional condition obtained higher scores for recall, transfer, hypotheses making, and total criterion. As with the results of the analyses for anxiety, these comparisons are significant for recall, hypotheses making, and total criterion. In contrast with the results for anxiety, none of the main effects for exhibitionism on the four criterion test scores is significant.

Interactions of Exhibitionism and Instructional Condition

None of the interactions of ES and instructional condition on the four criterion test scores is significant. The lines for the interaction of ES and instructional condition on transfer are parallel (no tendency toward interaction), whereas those for recall, hypothesis making and total criterion suggest interactions—in a direction opposite to prediction. This interpretation is unwarranted in view of the fact that the differences between the criterion test means for high and low ES within instructional conditions are very small. The largest of these differences is for total criterion. The difference between total criterion test means for high and low ES in the conventional condition is .6 (21.2—20.6). The comparable difference for high and low ES in the programed condition is .5 (18.4—17.9). As with the analyses for anxiety, it is again evident that there is no tendency (significant or otherwise) for the data to support the hypothesis concerning the interaction of exhibitionism and instructional condition on achievement.

Summary of Findings

For anxiety, 32 analyses, each involving an appropriate test of the hypothesis predicting an interaction between anxiety and instructional condition on achievement were accomplished. Four analyses appropriate to testing the hypothesis predicting an interaction between exhibitionism and instructional condition on achievement were accomplished. In all cases, the findings with respect to these hypotheses were inconclusive.

---

[5] None of the double interactions were significant.

Secondary findings from the analyses for this aspect of the research are summarized as follows:

1. The main effect of sex on achievement was nonsignificant in all analyses for anxiety and exhibitionism.
2. The main effect for mental age on achievement was significant in all analyses for anxiety and exhibitionism.
3. The main effect for instructional condition was significant for recall, hypotheses making, and total criterion in all analyses for anxiety and exhibitionism. These results indicate that the conventional mode was generally superior to the programed mode in teaching the learning material used in the research.
4. Anxiety generally had a negative effect on achievement in both instructional conditions. Significant main effects of anxiety on achievement are as follows:
    a. Generalized school anxiety on transfer.
    b. General classroom anxiety on recall, transfer, and total criterion.
    c. Parental pressure for achievement on hypotheses making and total criterion.
    d. Test anxiety on transfer.
    e. Total anxiety on transfer and recall.
5. None of the main effects for exhibitionism on the four criterion test scores was significant.
6. None of the additional double interactions included in the analyses (but not reported) was significant.

The practical significance of the findings summarized above may be ascertained from the differences between the means for each of the main effects. The data show that the differences between the means for each of the main effects of sex, mental age, and personality, even when significant, are generally small.

INSTRUCTIONAL MODE, COMPULSIVITY AND ACHIEVEMENT

Lawrence Wightman, Cornell University

Introduction

My task in the Learning Structures Project was to develop a paper and pencil self-assessment of compulsivity for eighth graders, for use in *testing* specifically the following null hypothesis: Mode of instruction (programed (P), conventional (C)) has no effect on the relationship between compulsivity and achievement. We expected that, when compared to the non-compulsive student, the compulsive student would do better with programed instruction than with conventional teacher instruction. This paper summarizes the steps and some of the data of the developmental procedure, describes the final test(s), and summarizes the results of the experimental year.

Test Construction

After a review of the literature relating to compulsivity, a definition was developed, composed of logical subcategories with several short descriptive phrases for each. About 240 items, representing the nine categories, were constructed and subjected to factor analytic procedures in two stages, during the developmental year.

Seven logically and empirically derived factors were found, similar to the original nine factors, but distributed in two general factors, one constructive (composed of meticulousness, tendency to finish, intolerance of incompleteness, and cautiousness) and the other unconstructive (composed of paralyzed initiative, intolerance of indefiniteness, and uncomfortableness in social relationships). The same items for the experimental year sample ($N > 1000$) were subjected to the same analysis and the results were identical.[1]

Although conventional item analysis techniques are not appropriae (see Cattell and Tsujioka, 1964), "Kuder-Richardson Formula 20" reliabilities, and item discriminations are somehow more familiar landmarks. As could be expected, the KR20 split-half reliabilities and the item discrimination indices are moderate for relatively heterogeneous tests. Constructive compulsivity has a KR20 reliability of .81 and a mean (upper half/lower half) item discrimination of 25.1. Unconstructive compulsivity has a KR20 reliability of .57 and a mean (upper half/lower half) item discrimination of 23.5. The subtest reliabilities ranged from .68 for meticulousness to −.07 for intolerance of indefiniteness. The subest item discriminations ranged from 37.6 for tendency to finish, to 29.1 for meticulousness.

Results

The analysis of variance involved four two-level classifications: sex, mental age (high, low), learning condition (P, C), and personality variable (high, low). There was a separate analysis for each type of compulsivity for each of the four achievement subtests: recall, transfer, hypothesis making, and total. (The statistical analysis has been covered in detail elsewhere in these proceedings. See R. E. Ripple, "Overview and Development of Instructional Materials.")

Main Effects

High constructive compulsivity is associated consistently with higher scores, although in no single case is there a significant difference. The conventional teaching obtained better results than the program, especially on simple recall (C mean = 8.2, P mean = 6.2; $p < .005$), to a lesser degree on hypothesis making (C mean = 4.28, P mean = 3.87; $.01 > p > .005$), and only slightly on transfer (C mean = 8.31, P mean = 8.05; n.s.). Maximum scores are 12, 6, and 13 respectively. Mental age differences, of course, are significant ($p < .005$) on all three achievement subtests and the total.

Interaction

Although some personality/condition interaction is present, it does not approach significance and, more importantly, it is opposite to what was expected. Students high in (constructive) compulsivity do relatively better in the conventional situation (e.g. on recall—P mean = 6.27, C mean = 8.38; low compulsivity—P mean = 6.16, C mean = 8.06).

---

[1] There is not space here to go into these data but they are available upon request. Also available: original categories with descriptive phrases, the final 62 items divided into the seven subtests, means and variances of all variables with correlation table, analyses of variance tables, split-half reliabilities on subtests, subtest item discrimination indices, and list of references.

Unconstructive compulsivity did not affect the student in the conventional classroom situation, and, after the programed instruction, those high did only slightly less well than those low. Instructional condition main effect was of the same importance as in the constructive compulsivity analyses.

Discussion

When a researcher fails to reject the null hypothesis, he should reconsider one or more of the following points: (1) the measuring instruments; (2) the research design (including treatment conditions); (3) the analysis technique; and/or (4) the rationale. Weakness in any one of the first three points can lead to failure to reject the null, when in fact, the null should be rejected. If, in truth, there are *no* differences it can not be proved. *If* careful consideration of the first three points shows no weaknesses which would lower sensitivity of the experiment, then the researcher *may* persuade himself to change his rationale and try again.

A full consideration of the *measuring instruments* will have to await the final report of the project. I feel that there is logical validity and some empirical validity for them but much more needs to be done with construct validation. This and other data such as test-retest reliabilities are needed before we can have complete confidence in the measuring instruments.

There are some unusual features of the *instruction* which should be mentioned. The most outstanding was that there was no homework requirement for either group. The tendency to organize one's time and finish things obviously would make little difference under such circumstances. The program was not given to the student to do at will, but rather administered during scheduled periods. No matter how hard it might be to get started or continue *on his own,* the student simply went to class and followed instructions.

This particular program was not constructed specifically for the truly compulsive child. He is sometimes asked to guess at answers, and in general, it is not a spoon-fed, only-one-possible-answer type of program. Therefore, there was probably little difference, for the compulsive student, from the regular, moderately structured classroom. In addition, the teachers of the conventional classes were supplied with excellent lesson plans, plus audiovisual materials, and undoubtedly only a few teachers would have put together, on their own, the kinds of lessons which resulted.

The *analysis of variance* was a relatively large, four factor design; school was used as the unit of analysis, and all the variables were dichotomized. Dichotomization doesn't make use of the full power of the data. In addition, to fill cells of analysis for each school, dichotomization was carried out separately for each school. Although two classes were matched for mental age *within* each school, the means varied considerably *between* schools so that some children classified as high M.A. in one school were actually lower than children classified as low in another school. Therefore, the achievement means for the high mental age groups (or high anxiety groups, or high compulsivity groups, etc.) for each school are unnecessarily variable, reducing the overall high mental age (and other) achievement means, raising the overall low mental age (and other) means and thereby reducing the mean differences.

These variations also lead to an amount of variance between schools that expands the error term, making the F tests less sensitive. They also contribute to the school-with-main-effects interactions, and the school-with-interactions, which were combined to make the error term.

All this said, it may still be argued that if practical, potent differences exist, even a "less sensitive" experiment would point them up. If one believes this, then compulsivity can be dropped from further consideration as an instructional variable. I feel that these data are specific to this program, conventional condition, and subject matter. Our experience has shown that treatment conditions need to be analyzed and specified explicitly, if not measured and perhaps specifically manipulated, in future attempts to obtain positive results that will contribute to an explanation of achievement.

INSTRUCTIONAL MODE, ACHIEVEMENT AND CREATIVITY

John S. Dacey, Boston College

The development of a battery of tests of verbal creativity and of convergent thinking is discussed first in this paper, followed by a discussion of the results of the experiment as they pertain to verbal creativity and convergent thinking. In constructing the test instruments, a survey of the literature was first undertaken. On the basis of this, it was decided that two basic aspects of verbal creativity should be operationally defined and measured: verbal creativity considered as a cognitive ability, and as a personality characteristic. A battery of 17 tests was constructed, administered to 300 eighth graders, and the results factor analyzed. A battery half as long as the original battery was selected. The new battery was administered to 300 eighth graders, and the results factor analyzed. A final battery one-fourth as long as the original set of tests was selected on the basis of these results. The measure of convergent thinking was constructed by the same procedure.

Of particular interest in this phase of the research was the discovery, by means of a varimax rotation of the first principle components analysis, that the loadings of Torrance's four creativity factors (fluency, flexibility, originality and elaboration) were grouped by *subtest* type, rather than by *factor*. Theoretically, score on these factors should not depend on the subtest used. This difficulty was corrected by scoring each of Torrance's tests only for that factor which loaded highest on it.

The major question asked by the Learning Structures Project was whether or not there is an interaction between type of instruction and the student's verbal creativity performance, on his score on each of the criterion tests. For example, do students high in originality perform better on the criterion tests when instructed by the conventional, as opposed to the programed, teaching technique? No such interactions were found. On the other hand, a number of interesting main effects of verbal creativity, as well as of mental ability and treatment conditions, on performance on the criterion test were found.

EVALUATION OF READING PROGRESS OF FOURTH GRADE PUPILS ENROLLED IN THE REMEDIAL READING PROGRAM OF THE AFTERSCHOOL STUDY CENTERS[1]

George Forlano & Jack Abramson, New York City Public Schools

The Purpose

The Afterschool Study Center Program is one of the steps taken by the Board of Education to compensate, as far as is possible within the operation of a school system, for the lack of maximum opportunity to learn academic skills as these skills are influenced by social and economic conditions. The objectives of this program are to provide remedial and other services beyond the regular school day program and to make available personnel, space, opportunity, and incentive for academic improvement.

Participants in the program have observed several indications, largely subjective in nature, which lead to the conclusion that the program is making positive contributions to the educational and emotional needs of many of the pupils involved. The Bureau of Educational Research has now proceeded to collect objective data concerning the effectiveness of selected aspects of the program. The area of reading was selected for study because reading activities comprise a very large part of the Afterschool Study Center Program and reading skills are the core of elementary education.

The Problem

Data were collected to seek answers to the following specific questions:

1. Do fourth grade pupils enrolled in the remedial reading program of the Afterschool Study Centers show greater gains in reading achievement over a period of a year than comparable fourth grade pupils in the same schools who had not enrolled in the Afterschool Study Center Program?
2. At different levels of reading retardation, do fourth grade pupils enrolled in the remedial reading program of the Afterschool Study Centers show greater gains in reading achievement over a year than do comparable fourth grade pupils in the same school who had not enrolled in the Program?
3. How does the percent of the Afterschool Study Center fourth grade pupils who improved one school year or more in reading compare with that of comparable fourth grade pupils not enrolled in the remedial program?
4. What is the relationship between attendance in the program and reading achievement among the fourth grade pupils enrolled in the remedial reading program of Afterschool Study Centers?

The Pupils

The experimental group was drawn from elementary schools having After Study Centers and consisted of 4th grade pupils enrolled in the remedial program as of October 1964 and who had been in the same school as third graders in April 1964. Since a goal of the program was to emphasize reading remediation, all experimental pupils selected were restricted to those who tested at grade placement or below. This meant that these experimental pupils achieved a reading grade equivalent of 3.7 or below as of April 1964 when the children were in the third grade.

The control group consisted of pupils drawn from the same school as their experimental counterparts. Each pair was matched on the following criteria: grade status, sex, chronological age within plus or minus three calendar months, and reading comprehension grade equivalent within plus or minus three school months.

On the basis of the above matching criteria 1,521 matched pairs were obtained from 145 schools in 4 boroughs where Afterschool Study Centers were in operation.

The Tests Used

All pupils included in the study were administered the Metropolitan Upper Primary Test, Form C, in April 1964 as part of the citywide third grade testing program, and the Metropolitan Upper Primary Reading Test, Form A or the Metropolitan Elementary

---

[1] We appreciate the cooperation given us by Assistant Superintendent J. Wayne Wrightstone of the Office of Educational Research. (New York City Board of Education). A copy of the full report may be obtained from the Bureau of Educational Research.

Reading Test, Form A in April 1965 as part of the fourth grade citywide testing program.

Summary of Findings

The major findings are as follows:

1. Fourth-grade pupils enrolled in the remedial reading program of the Afterschool Study Centers show a significantly greater gain in reading achievement of approximately two school months more than comparable fourth-grade pupils in the same schools who had not enrolled in the program. Over the period of one school year the experimental pupils, as a group, grew 9.8 school months as compared with a growth of 7.7 school months for comparable control pupils.
2. At each of 3 levels of reading retardation, determined at initial test time in the third grade, fourth-grade pupils enrolled in the remedial reading program of the Afterschool Study Centers again showed significantly greater gains in reading achievement (approximately two school months) than comparable fourth-grade pupils in the same school who had not enrolled in the program.
3. A significantly larger percentage (approximately 15 percent) of the fourth-grade pupils enrolled in the remedial reading program of the Afterschool Study Centers improved one school year or more in reading when compared to fourth-grade pupils not enrolled in the program.
4. Fourth-grade pupils enrolled in the remedial reading program of the Afterschool Study Centers and who attended no less than 55 instructional sessions showed significantly greater gains in reading achievement (approximately one and a half school months) than did those fourth-grade pupils who attended no more than 27 instructional sessions of the Afterschool Study Centers. In general there was a positive association between attendance in the Afterschool program and reading achievement for these fourth-grade pupils.
5. Experimental pupils in the upper group in attendance when compared to their control counterparts showed an average gain of about 3 school months over the controls; experimental pupils in the lower group in attendance gained 2 school months over their control peers.

In general, the effect of enrollment in the remedial program in reading on the participating fourth-grade pupils was to increase their rate of progress in reading as compared to that of the pupils who were enrolled in the Afterschool Study Center Program. On the average the experimental pupils as a group gained about 27 percent more than the control pupils. Largest gains were observed for the subgroups of experimental pupils who were most retarded in reading and who attended more regularly. A positive relationship appeared between the amount of attendance in instructional sessions and growth in reading achievement.

In summary, the experimental group maintained its growth over the period studied while the control group became more retarded in reading in relation to grade placement. In general, one of the major objectives of the program—the reading improvement of the participants—was realized.[2]

---

[2] In a subsequent study, not reported here, the experimental pupils were compared with control pupils in schools with no Afterschool Study Centers. Positive results for the experimental pupils were again observed.

BEGINNING READING—THE EFFECTIVENESS OF i.t.a. AND T.O.[1]

Harold J. Tanyzer, Harvey Alpert & Lenore Sandel, Hofstra University

Problem and Objectives

This investigation, which was initiated during the 1964-65 school year, is presently in its third year, and is planned to continue until all children in the various treatment groups have completed third grade. This report is based upon the research of the first two years of the study.

The purposes of this study were: (1) to compare the effects of different orthographies—traditional orthography (T.O.) and the Initial Teaching Alphabet (i.t.a.)—on the reading and spelling achievement of children who were introduced to formal reading instruction either in kindergarten or first grade, in i.t.a. or T.O.; (2) to evaluate the effect of introducing formal reading instruction at the kindergarten level; and (3) to determine the longitudinal effects of teaching reading in i.t.a. and T.O. at the end of the first, second, and third grades.

Procedure

Treatments

During the first phase of this study, each of eleven Long Island school systems contributed from two to eight kindergarten and first-grade classes. The original sample included 1,168 kindergarten children in 44 classes and 852 children in 34 first-grade classes. Reading instruction was initiated with four groups: two groups, one at the kindergarten level and the other at the first-grade level, were taught to read by the i.t.a. medium; and two groups, also at the kindergarten and first-grade levels, were instructed in T.O.

Groups learning to read in T.O. were taught by varying methods, depending upon the teachers' preferences. Teachers were allowed to utilize procedures that were generally approved by language arts specialists and to instruct in a basal reader, or multibasal approach, or individualized reading approach.

The i.t.a. classes were taught to read in a series entitled the *Early-to-Read i/t/a Program,* a set of readers specially constructed to exploit the advantages of i.t.a.'s phonemic regularity.

Teacher Selection and Training

Within each district, principals were requested to recruit teachers who would be willing to volunteer for the study with no preconceived notion whether they would teach in the i.t.a. or T.O. medium. From those teachers volunteering for the study, principals were instructed to match teachers on the basis of teaching experience, training, and teaching competence. Teachers attended a preschool three-day workshop designed to provide a theoretical basis and practical application of approved methods of teaching reading in T.O. and i.t.a. Following the workshop sessions, teachers were assigned randomly to i.t.a. or T.O. classes. In the second year of the study, the first-grade teachers participating during the previous year were assigned to work with the new first-grade classes formed from the previous year's kindergarten population. Second-grade teachers were selected and trained by procedures similar to those used with the previous year's first-grade teachers and then were assigned to one of the two experimental groups on a random basis.

At the conclusion of the first phase of the study, each first-grade i.t.a. and T.O. class was moved to second grade as a unit. The kindergarten groups were pooled within their medium of instruction and assigned to first-grade classes.

Pretesting

The following readiness and intelligence tests were administered in September 1964 to the first-grade classes and in November to the kindergarten classes: *Metropolitan Readiness Tests* and the *Pintner-Cunningham Primary Test.*

Posttesting

At the completion of the first year of the study, the *Stanford Achievement Test,* Primary I was administered in traditional orthography to all children in the first-grade i.t.a. and T.O. classes. No measure was administered to the kindergarten groups. At the completion of the second year of the study, the appropriate battery of the *Stanford Achievement Test* was administered to the first- and second-grade groups.

Statistical Analysis

Both analysis of variance and covariance procedures were employed.

---

[1] The work upon which this report is based was supported jointly by The Education Council for School Research and Development, Hofstra University, and the New York State Education Department, under Article 73, section 3602a, subdivision 14 of the State Education Law.

Results and Conclusions

1. Introducing i.t.a. to kindergarten children in a formal reading program does result in significantly better word recognition and word analysis than that attained by children who learn in T.O. in kindergarten, when both groups of children are measured at the end of first grade. No significant difference is found in comprehension, but the T.O. group is significantly better in spelling.
2. Introducing i.t.a. to kindergarten children does not result in significantly better reading and spelling achievement than that attained by children who began reading instruction in first grade in i.t.a., when both groups are measured at the end of first grade.
3. Introducing reading in T.O. to kindergarten children does not result in significantly better reading and spelling achievement than that attained by children who were introduced to reading in T.O. in first grade, when both groups are measured at the end of first grade.
4. Introducing reading in i.t.a. to kindergarten children results in significantly higher word recognition than that attained by children who were introduced to reading in first grade in T.O., when both groups are measured at the end of first grade. No significant differences are found in comprehension or word analysis, but the T.O. group was significantly better in spelling.
5. Introducing i.t.a. to first-grade children results in significantly better word recognition and word analysis than that attained by children who began reading instruction in kindergarten in T.O., when both groups are measured at the end of first grade. No significant difference is found in comprehension, but the T.O. group is significantly better in spelling.
6. Introducing reading to first-grade children in i.t.a. does not result in significantly better reading achievement, and does result in significantly poorer spelling achievement than that attained by children who learned T.O. in first grade, when both groups are measured at the end of first grade. At the end of second grade the i.t.a. group is significantly better in word meaning, word study skills, and spelling achievement. No significant difference is found in comprehension.

AN EVALUATION OF THE INITIAL TEACHING ALPHABET IN THE FIRST GRADES OF THE CANANDAIGUA ELEMENTARY SCHOOL

Marion Walsh, Canandaigua Public Schools

Phonetic and augmented alphabets have been suggested and tried in elementary education for more than a century. While educators have recognized the complexities of our language and the difficulties in teaching it, these early alphabets received no general acceptance.

During the past 5 years, however, more and more people have become interested in a new augmented alphabet being introduced and researched in England. This alphabet developed by Sir James Pitman and now termed The Initial Teaching Alphabet, "I.T.A.," produced encouraging results in early experimentation. These experiments in British schools encouraged educators in the United States to develop experimental programs using the same teaching materials.

For the past several years there has been a growing feeling among many of the staff members at Canandaigua that further exploration into the Initial Teaching Alphabet was worth consideration.

After exploration there still remained questions as to how well this Initial Teaching Alphabet would work in the Canandaigua schools. Also, there appeared to be little data available on achievement of children beginning reading using I.T.A., compared with children beginning reading using the approaches suggested in a basal reading series, generally referred to as "Traditional Orthography" T.O. This study was developed in the hope of answering questions in both of these areas.

Procedures

At a meeting with the remedial reading teachers and administrative staff, the decision was made that we should request permission and funds from the Superintendent of Schools and the Board of Education to conduct two experimental groups with the same number of children in the control situation for 1965. This was approved, but later, for several reasons, it was decided to double the size of both groups.

Students

The I.T.A. experimental group was originally composed of 100 children whose parents had indicated an interest for them to participate in the program. A check of teachers' evaluations and other data revealed that this group was made up of higher achievers than the group of approximately 180 students who were not selected. To partially counteract this problem, a stratified random sample of 100 students was drawn from the group of 180 students to form the T.O. control group. This stratified sample was made according to teacher ratings of students' capabilities for success in grade one.

Chart No. 1—shows the children assigned to the Initial Teaching Alphabet program and those chosen for the control situations under the heading of Teacher Judgement on an A. B. C. D. E. basis.

Teacher Judgement

Comparisons of the I.T.A. and the T.O. groups on the Metropolitan Readiness Test and on the Bender Gestalt indicated that the groups were not significantly different on either of these measures.

I. T. A.	T. O.
A + B rating — 49	A + B rating — 50
C rating — 32	C rating — 35
D + E rating — 15	D + E rating — 13

The I.T.A. group was taught in four sections while the T.O. group was randomly assigned to the remaining seven sections in grade one. Teachers working with students in the T.O. group were not informed as to who these students were. Thus, they could not identify the T.O. group students in their classes.

Teachers

The determination of the teachers for the Initial Teaching Alphabet groups by the administration was made on the basis of interest and enthusiasm shown by several teachers toward the program. In the final selection, 2 kindergarten teachers were chosen—one having had many years of experience in both kindergarten and first grade and the other a teacher of 2½ years experience. Two first grade teachers were chosen, one having had 2 years experience in first grade and the other having had 8 years experience in first grade. It might be well to note here that our 7 teachers of the T.O. groups were not significantly different in their teaching experience. Their experience varied from one teacher with no previous teaching experience to one with many years of experience in kindergarten and first grade. The other 5 teachers had, on the average, approximately 6 years of teaching experience.

Program

The study was begun in September of 1965 and continued through June, 1966. The T.O. group received the program that had been used for several years, The Ginn Basic Reading Series. The I.T.A. group used The Early To Read Series, plus I.T.A. materials specially prepared by the participating teachers. The major differences that occurred in the school program were: (1) Many more meetings of the teachers of the I.T.A. group (both formal and informal); (2) General information meetings for all primary teachers.

An informal evaluation of the program was done in January, with formal achievement testing during May and June in keeping with regular school schedules.

Analysis of Data

In May of the school year 1965-66 the Stanford Achievement Test, Primary I Battery, Form W was administered to all of the first graders. Those in the T.O. groups were given the traditional edition, those in the I.T.A. groups were given the I.T.A. version.

Tables I and II

Certain comments seem to be in order:

A. The significant difference of 2.0 to 2.8 in Word Reading was not startling due to the fact that this is a conclusion we had already drawn as I.T.A. seemingly gives children the necessary skills to attack new words.

B. The small differences in paragraph meaning and vocabulary were certainly below our expectations for it was our feeling on the basis of previous predictions that paragraph meaning for the I.T.A. people would be considerably higher. All the teachers using the I.T.A. medium of instruction and the T.O. teachers were somewhat perturbed by the relative low rating of both groups in paragraph meaning and vocabulary. The thought was expressed that this might have been due somewhat to the format of the test as it was considerably different than the format of tests being presented to our children heretofore or materials to which they had been exposed in their Activity Books.

C. It is well to note the difference in the means of the T.O. people (2.7) and the I.T.A. people (3.8) in Word Study Skills. It is obvious that the children of the I.T.A. groups showed greater knowledge in audio perception of beginning and ending sounds. Children in the I.T.A. groups were much more alert in the areas of phonics and phonograms.

D. Before we started our I.T.A. program there was a great deal of concern on the part of all of our teachers and parents as to how well our children might do in the area of spelling. This particular test shows that there should be no real concern in this area as the I.T.A. people outranked those in the control groups on the first testing, when the I.T.A. people were using their own medium of spelling. There was probably no difference when they were spelling the same words in T.O.

E. It should be noted that there is little difference between the I.T.A. people and the T.O. people in the area of arithmetic. This is indicative of the fact that the I.T.A. children were not neglected in this area even though a great deal of emphasis was placed upon the reading program for the school year 1965-66.

Discussion

As the school year closed, it became obvious that approximately 80 percent of the children had made a complete transfer from the I.T.A. medium to the T.O. and that they were able to handle material in traditional orthography with ease. Further evidence of this transition comes from the testing of the I.T.A. children on the Stanford Achievement Test using the T.O. version. It is interesting to note that there is very little difference in some of these areas, with the greatest difference being shown in the area of spelling.

We are encouraged from what we can see at this point for we find children enthused about reading and writing creatively. There appear to have been fewer frustrations in the area of the beginning reading than we have seen in the past. Our teachers are convinced that this approach to beginning reading is sound and it is a medium through which they can use many of the techniques that they have used heretofore for effective teaching.

While results are very encouraging at this point, certain questions and problem areas must be considered—the biggest one being:

What type of program should be provided for the I.T.A. children moving into the second grade for the school year 1966-67 which will keep within reasonable limits and still meet the needs of these boys and girls without destroying our control situations?

In summary, it might be said that we are encouraged and nearly convinced, but longitudinal studies will still be needed before making final decisions on I.T.A.

RELATIONSHIP OF READING ACHIEVEMENT AND MOTIVATION

Irving Zweibelson, New Rochelle Public Schools & Francis J. Lodato, Manhattan College

Several observations are worthy of note concerning reading achievement. Good readers are generally the product of homes which maintain a high interest in reading. Poor readers frequently come from homes which maintain a lower interest in reading. For example, Sheldon and Carillo (1952) indicated that the number of books in the home was related to the percentage of good and poor readers. Milner (1951) found that the reading ability of grade one children was related to family social status, verbal family environment and the emotional family interaction. The findings mentioned above and countless other studies suggest that reading might well be related to, and perhaps a function of, personality. Reading ability may depend as well on attitudes developed within the learning environments children are exposed to in home and school. Certainly, reading itself and learning attitudes must be greatly influenced by home and school experiences. In the light of the attention focused on preschool programs, it seems logical and timely to investigate the likelihood that a student's attitude toward and interest in reading is likely to be related to his reading achievement and skill. It seems clear that a student's motivation to achieve in school can be affected by his success or failure in school. His desire to achieve in a satisfactory manner probably will affect his motivation to read. Therefore, early attention to reading can be meaningless if learning attitudes become negative because of later poor school and home experiences.

Reading, when compared with other elementary school subjects, seems to provide a particularly sensitive barometer of a pupil's motivation to perform well in school. One reason might well be the key place reading holds as an educational tool. This does not necessarily mean, however, that a disinterested reader is a poorly motivated child. The processes affecting reading and motivation are too complex to allow so simple an assumption. It does seem probable that knowledge of the child's motivation to achieve is likely to be a useful index in providing help to improve school performance, as well as reading performance.

The present study undertook an exploratory evaluation of the "motivation level" of a small sample of students for comparison with the available scores on a standard school achievement test. The population tested was the entire 5th grade, 3 classes, of a single school. Sixty-five students with an I.Q. range of 84 to 136, from a middle class racially integrated neighborhood, were given a motivation inventory by the teachers who read 80 questions to their respective classes.

The major purpose of the study was to attempt to assess the motivation of 5th grade students, as well as the relationship of reading achievement to motivation. It was hypothesized that 5th grade elementary school students were sufficiently sophisticated to respond to a motivation inventory, developed by Frymier (1962). It was expected that the motivation index would have a significant and positive relationship with reading achievement test scores. The results of a correlation analysis, shown in Table I, provide support for the hypothesis that the pupils' scores on the motivation index were related to their grade equivalent reading test scores. It was also true that there was a similar relationship of the motivation scores with spelling and language scores. In the case of reading and language study skills the correlations of .28 and .25 respectively were significant at the .05 level. The highest of the correlations, shown in Table I was .38 (significant at the .01 level) for the language test and motivation. It should be particularly noted that I.Q. was not significantly correlated with motivation (R = .10). The correlations of motivation and word knowledge, arithmetic computation, arithmetic problem solving and study skills, were not significant.

These findings provide support for the assumption that reading and language are interacting factors affecting students motivation and general achievement. The teacher and the reading specialist, using pupil response inventories, can assess the attitudes and interests of pupils for the purpose of encouraging changes in the personal attitudes or the instructional process. Such information can be useful not only for determining how students feel about school but also for exploring further the possibility that motivation and reading might be considered as *basic* factors affecting the learning process.

TABLE I

MEANS, STANDARD DEVIATIONS, AND CORRELATIONS OF I.Q. AND ACHIEVEMENT TEST SCORES WITH MOTIVATION INDICES

(N = 65)

Test	$\overline{X}$	S.D.	R
Motivation	102.5	22.	
I.Q.	120.0	10.	.10
Word Knowledge	6.97	.50	.17
Reading	6.08	.32	.28[1]
Spelling	6.88	.41	.38[2]
Language Study Skills	7.58	.47	.25[1]
Arithmetic Computation	6.65	.65	.06
Arithmetic Problem Solving	6.86	.54	.15
Study Skills	6.12	.50	.14

[1] Significant at .05 level
[2] Significant at .01 level

List of Participants

LIST OF PARTICIPANTS

Abramson, David A., Bureau of Curriculum Research, New York City
Abramson, Jacob, 719 Frisio Ave., Far Rockaway
Adler, Leona K., Hunter College, New York City
Alland, Peter W., Guilderland Central Schools, Guilderland
Allen, Joseph E., Plattsburgh City School District, Plattsburgh
Andrews, Gloria, University of Rochester, Rochester
Armstrong, Charles, New York State Education Department, Albany
Aronow, Miriam S., Long Island University, Brooklyn
August, Joel A., University of Bridgeport, Bridgeport, Conn.
Bache-Wing, Carl W., Maine Education Department, State Office Bldg., Augusta, Maine
Bagley, Clarence H., SUC Cortland
Baker, Elizabeth, New York State Education Department, Albany
Barabasz, Arreed F., SUNY at Buffalo
Barnes, James R., New York State Education Department, Albany
Barnhard, Ivan H., 85 Blauvelt Ave., West Haverstraw
Barz, Anita I., St. John's University, Jamaica
Beach, Mark A., Cornell University, Ithaca
Becker, Norman W., Cornell University, Ithaca
Behan, G. Wallace, Middletown High School, 162 Wickham Ave., Middletown
Benson, Gregory M., Averill Park Central School, Averill Park
Berkowitz, Howard, SUC Oneonta
Berkum, Howard D., 1314 3rd St., Rensselaer
Bernardo, Charles, Columbia University, New York City
Bidwell, Wilma W., SUNY at Albany
Bird, Charles A., Fordham University, New York City
Blackwood, Ralph O., University of Bridgeport, Bridgeport, Conn.
Bligh, Harold F., Harcourt, Brace & World, Inc., 757 Third Ave., New York City
Blixt, Sonya Lynn, SUNY at Albany
Bloom, Benjamin S., University of Chicago, Chicago, Ill.
Boaz, Ruth L., N.Y.S. Library Development Division, 40 Steuben St., Albany
Boquist, Robert D., SUNY at Albany
Boyce, Jean, SUC Fredonia
Boyle, John F., Rev., Fordham University, Stoney St., Shrub Oak
Brady, James J., New York State Education Department, Albany
Braun, John R., University of Bridgeport, Bridgeport, Conn.
Bristow, William H., Bureau of Curriculum Development, 130 West 55th St., New York City
Broadbent, Frank W., SUC Brockport
Brod, Pearl, 64 Beach Ave., Larchmont
Brown, Dorothy O., 4329 Onondago Blvd., Syracuse
Brown, Foster L., SUC Oneonta
Brown, Janet R., Queens College, Flushing
Brown, Roscoe C. Jr., New York University, New York City
Bruce, Bernstein H., Brooklyn Academy, Brooklyn
Bulles, Berton, 45 Church St., Nassau
Burdick, Robert E., West Genesee Central School, Camillus
Callahan, William T., The Education Council, 131 Mineola Blvd., Mineola
Calvert, John F., BOCES, 5203 West Genesee St., Camillus
Campbell, Bertha D., New York State Education Department, Albany
Campbell, William J., Hicksville Public Schools, District #17, Hicksville
Carter, Donald E., SUNY at Buffalo
Castiglione, Lawrence V., Queens College, Flushing
Castillo, Jovito B., University of Rochester, Rochester
Chapline, Elaine, 567 Ridgeland Terr., Englewood, N.J.
Clark, Richard M., SUNY at Albany
Cleminson, Geraldine, St. John's University, New York City
Clifford, Peter C., SUNY at Delhi
Conlin, Reverend Matthew, Siena College, Loudonville
Connelly, John E., SUC Fredonia
Cons, Ruthann G., Bethlehem Central Schools, Delmar
Cosman, Charles B., Islip Public Schools, 215 Main St., Islip
Coxe, Warren, 120 Salisbury Rd., Delmar
Crane, Edmund, New York State Education Department, Albany
Crew, John L., Maryland State Dept. of Education, Baltimore, Md.
Cuddy, Brother Henry, St. Francis College, Brooklyn
Cullinan, Paul A., New York University, New York City
Dacey, John, Boston College, Chestnut Hill, Mass.
Danowski, Charles E., Columbia University, New York City
Darlington, R. P., Cornell University, Ithaca
Dean, Jeff A., SUC Oneonta
De La Fleur, Frederick, N.Y.S. Schools Boards Association, 111 Washington Ave., Albany
Delany, John, SUC Oneonta
Denny, David A., SUNY at Oneonta
Dentler, Robert, Center of Urban Education, 33 West 42d St., New York City
Dever, Robert W., Mineola Public Schools, Mineola
Di Lorenzo, Louis T., New York State Education Department, Albany
Dodd, John M., SUNY at Buffalo
Dominy, Mildred M., SUC Plattsburgh
Donely, Donald T., SUNY at Albany
Doyle, Esther C., 27 Snowden Ave., Delmar
Droessler, Earl G., SUNY at Albany
Drumheller, Sidney J., SUC Brockport
Dubnick, Lester, U.F.S.D. #3, 37 Panorama Dr., Huntington
Ducanis, Alex J., SUNY at Binghamton
Duffy, Gerald G., SUC Fredonia
Duken, Gilbert A., Plattsburgh School Districts, Plattsburgh
Dwyer, Thomas, New York State Education Department, Albany
Eagle, Norman, 412 Tenafly Rd., Englewood, N.J.
Econome, Madelyn B., Scotia-Glenville Central Schools, Scotia
Egner, Joan Roos, Cornell University, Ithaca
Essex, Stewart R., Rhode Island Dept. of Education, 228 Spring St., East Greenwich, R.I.
Ether, John A., SUNY at Albany
Farr, S. David, SUNY at Buffalo
Finger, John A., Rhode Island College, Providence, R.I.
Flurry, Ruth C., New York State Education Department, Albany

Fontes, Patricia J., Boston College, Boston, Mass.
Forman, Joseph, New York State Education Department, Albany
Formanek, John J., New Lincoln School, New York City
Formanek, Ruth, Hofstra University, Hempstead
Forte, Joseph, New York State Education Department, Albany
Foster, Helen, SUC Geneseo
Fox, David J., The City College of New York, New York City
French, L. Ronald, SUC Oneonta

Gallo, Sue, SUC Oneonta
Garrow, Robert M., Board of Education, Plattsburgh
Gaverick, Charles M., SUNY at Buffalo
Gioia, Vito, One North Elm St., Schenectady
Giragosian, Sandra J., Syracuse University, Syracuse
Glock, Marvin D., Cornell University, Ithaca
Gordon, Barbara, 48 Valley Circle, Rochester
Gordon, Leonard V., SUNY at Albany
Gordon, Sheila C., United Federation of Teachers, 300 Park Ave. South, New York City
Gorman, Sister Marie Helen, St. John's University, Jamaica
Greenham, William H., Brighton High School, Rochester
Grennell, Robert L., SUC Fredonia
Griffin, Mary D., Boston College, Boston, Mass.
Guenther, Richard C., SUNY at Buffalo

Hageny, William J., SUC New Paltz
Hall, Daniel R., SUC Geneseo
Hallwell, Joseph W., SUC Cortland
Halperin, Silas, Syracuse University, Syracuse
Halverson, Paul, Syracuse University, Syracuse
Hanna, Gerald J., Harcourt, Brace & World, Inc., Testing Department, 575 Third Ave., New York City
Harkness, Peter, New York State Education Department, Albany
Harmon, Richard P., Long Island University, Long Island
Harriger, James C., 4074 Candy Lane, Manilus, Syracuse
Harris, Albert J., Division of Teacher Education, City of New York, 345 East 80th St., New York City
Harris, Mary, New York State Education Department, Albany
Haskell, Eugene F., State of Maine, Department of Education, 30 Central St., Randolph, Maine
Hastings, James, SUC Oswego
Havens, John W., SUC Oneonta
Hayden, Robert R., SUNY at Albany
Hayes, Dorothy, SUC New Paltz
Hayward, Priscilla, New York State Education Department, Albany
Heller, Robert W., SUNY at Buffalo
Henderson, Edward H., New York University, New York City
Hicks, Jean E., SUC Fredonia
Higgins, Conwell, Giffen School, 274 So. Pearl St., Albany
Hill, G. Eugene, SUC Plattsburgh, Plattsburgh
Hoff, John F. Jr., 61 Parrott Rd., West Nyack
Hoffman, Hillard R., Pennsylvania Department Public Instruction, Harrisburg, Pa.
Hofman, Richard J., SUNY at Albany
Hogan, Gail F., SUNY at Albany
Horend, James L., Glenham Elementary School, Beacon
Hughes, Raymond E., Canandaigua Public Schools, Canandaigua
Hughes, Warren A., U.S. Office of Education, Washington, D.C.
Hunt, Rolfe L., International Journal of Religious Education, 475 Riverside Dr., New York City
Hunter, Delores M., 40 Emory St., Hempstead

Immegart, Glenn L., University of Rochester, Rochester

Iversen, Jack R., SUNY at Oneonta

Jackson, R. Peter, Cornell University, Ithaca
Jacobson, William, Pelham High School, Pelham
James, W. Raymond, SUC Oneonta
Jary, Paul W., Siena College, Loudonville
Jaspen, Nathan, New York University, New York City
Johnson, S. W., SUC Plattsburgh
Johnston, Paul H., Delaware Dept. of Public Inst., Dover, Del.
Joiner, Lee M., Hofstra University, Hempstead
Joslin, Paul H., Drake University, Des Moines, Iowa
Justman, Joseph, Center on Urban Education, 33 West 42d St., New York City

Karker, Helga, 69 South Lake Ave., Albany
Kaye, Mildred, Brooklyn College, Brooklyn
Kenny, J. Ruport, University of Rochester, Rochester
Kenyon, Raymond, Regional Director, Mid-Hudson Supplementary Education Center, New Paltz
King, Dorothy H., New York State Education Department, Albany
Kiser, Chester L., SUNY at Buffalo
Kline, Francis, 75 Lake St., Oshkosh, Wis.
Knapp, Elizabeth A., SUNY at Binghamton
Knapp, Thomas R., University of Rochester, Rochester
Korngold, Robert D., GEAR Associates Inc., 130 Cherry Ave., Delmar
Kosky, Elizabeth M., Yonkers Public Schools, Yonkers
Kraus, Philip E., Hunter College, New York City
Krebs, Robert E., Pace College, New York City
Kreisman, Leonard, New York State Education Department, Albany
Krevisky, Joseph, Center for Urban Education, 33 West 42d St., New York City
Kurland, [illegible]man, New York State Education Department, Albany

Laabs, Charles, Columbia University, New York City
LaMantia, Gerald P., Ridgefield High School, Ridgefield, Conn.
Lamie, Richard G., Valhalla Schools, 212 W. Post Rd., White Plains
Lancione, Antonio M., Plattsburgh City Schools, Plattsburgh
Lashinger, Donald R., SUNY at Albany
Lauver, Paul M., Pennsylvania Dept. of Public Instruction, Harrisburg, Pa.
League, James B., Maryland State Dept. of Education, Baltimore, Md.
Leese, Joseph, SUNY at Albany
Lehey, James, Baldwin Public Schools, Baldwin
Leibler, Seth N., Communicable Disease Center, Atlanta, Ga.
Leiner, Marvin, Queens College, Flushing
Lev, Joseph, New York State Education Department, Albany
Liu, B. Alfred, City University of New York, New York City
Livak, Frank H., Connecticut State Education Department, P.O. Box 2219, Hartford, Conn.
Lodato, Francis J., Manhattan College, New York City
Lohman, Maurice A., City University of New York, New York City
Long, Kathryn L., SUC Oneonta
Lonsdale, Richard C., New York University, New York City
Losee, Marshall, New York State Education Department, Albany
Lowe, W. T., Cornell University, Ithaca
Lucow, William H., Education Division, Dominion Bureau of Statistics, Government of Canada, Ottawa 3, Canada

Lutz, Frank W., New York University, New York City

MacQueen, Gordon A., Ramapo Central School District #2, Ramapo
Madaus, George F., Boston College, Boston, Mass.
Mahieu, George W., Bayshore Public Schools, 1415 E. Third Ave., Bayshore
Margolius, M. Kenneth, New York State Education Department, Albany
Martin, George M., Beacon Schools, 70 South Ave., Beacon
Mason, Jack L., Syracuse University, Syracuse
Matthews, Frederic, STACO, 80 Cornwell Ave., Williston Park, Long Island
McCabe, Donald P., Siena College, Loudonville
McCann, Richard A., U.S. Office of Education, Bureau of Research, Washington, D.C.
McColly, William, SUC Oswego
McFarland, Donald F., SUC Fredonia
McHugh, Paul F., Boston College, Boston, Mass.
McKay, Richard K., Maryland State Dept. of Education, Baltimore, Md.
McKelvey, Troy V., SUNY at Buffalo
McLoughlin, William P., St. John's University, Jamaica
McMorris, Robert F., SUNY at Albany
McNamara, James F., Department of Public Instruction School System, Harrisburg, Pa.
McTernan, John W., SUC Plattsburgh
Mehit, George, New York State Education Department, Albany
Meno, Lionel W., SUC Fredonia
Metty, Michael P., Syracuse University, Syracuse
Meyer, Donald, Syracuse University, Syracuse
Mildram, Barbara, New York State Education Department, Albany
Miller, Herbert R., Syracuse University, Syracuse
Millman, Jason, University of Illinois, Urbana, Ill.
Mitchell, James V., University of Rochester, Rochester
Momot, Arthur P., City School District, Plattsburgh
Morine, Greta G., Hofstra University, Hempstead
Morrision, Frances, SUNY at Albany
Morrison, J. Cayce, Park Crest Apt. G-3, 1700 South West Ave., Jackson, Mich.
Morrissey, Madeline M., Board of Education, 110 Livingston St., Brooklyn
Mother Mary Jeanne, Mt. Carmel, St. Benedicta School, White Place, Staten Island, New York
Movesian, George, Hunter College, New York City
Mulligan, Robert E., Fords Junior High School, Fords, N.J.
Murphy, Gerald, Central School District #3, Locust Valley
Murray, Kay, 100 LaSalle St., New York City
Murdoch, Robert L., SUNY at Albany
Musella, Donald F., SUC Cortland
Mustico, Thomas W., SUNY at Albany

Nasca, Don F., SUC Brockport
Navilia, Linda M., Junior College of Albany, Albany
Nelson, Carl B., SUC Cortland
Newbruger, Barbara N., University of Bridgeport, Bridgeport, Conn.
Nikoloff, Sayra B., Skidmore College, Saratoga Springs
Nuccio, Vincent C., Boston College, Chestnut Hill, Mass.

Oakes, Russell C., SUC Geneseo
O'Connell, William B., University of Rochester, Rochester
O'Hare, George J., 5 Francis Ave., Albany
Okes, Immogene E., U.S. Office of Education, Washington, D.C.
O'Neill, Patrick J. Rev., Catholic School Department, Fall River, Mass.
O'Reilly, Robert P., Cornell University, Ithaca
Orlando, Stephen, SUC Fredonia
Owens, Robert G., Brooklyn College, Brooklyn

Palermo, Richard D., SUC Oneonta
Panther, Edward E., SUNY at Buffalo
Passy, Robert A., New York State Education Department, Albany
Paulus, Dieter H., University of Connecticut, Storrs, Conn.
Peck, Melanie I., Syracuse University, Syracuse
Peterson, Raymond D., SUC Plattsburgh
Plumb, Cora S., NYS Schools Boards Association, 111 Washington Ave., Albany
Potts, Marion H., SUC Cortland
Powell, Leonard, New York State Education Department, Albany
Powers, James E., Syracuse University, Syracuse
Privateer, George J., St. Bonaventure University, St. Bonaventure
Pruzek, Robert M., SUNY at Albany
Pyle, Wilma J., SUC Fredonia

Quick, Custer R., Horace Mann School, Craig St., Schenectady
Quirolgico, Delfin G., UN Development Programme, 866 U.N. Plaza, New York City

Raub, J. Robert, University of Rochester, Rochester
Rector, Douglas F., SUC Fredonia
Reddy, Howard, Siena College, Loudonville
Reed, Horace, Skidmore College, Saratoga Springs
Reed, L. John, Skidmore College, Saratoga Springs
Reeling, Glenn E., Jersey City State College, N.J.
Reeve, Theo, New York State Education Department, Albany
Reimer, Charles T., SUC Oneonta
Reiner, William B., Hunter College, New York City
Reynolds, Richard J., SUNY at Albany
Ripple, Richard E., Cornell University, Ithaca
Rizzo, Paula, SUC Oneonta
Robbins, Harold, Queens College, Flushing
Roberts, Julian, Yeshiva University, New York City
Robertson, Alan H., New York State Education Department, Albany
Rock, William C., Rochester City School District, 13 Fitzburgh St. South, Rochester
Roden, Aubrey H., SUNY at Buffalo
Rosenbach, John, SUNY at Albany
Rosenstein, Joseph, Lexington School for the Deaf, 904 Lexington Ave., New York City
Rosett, Henry L., Institute of Psychiatry, 51 East 90th St., New York City
Ross, Paul C., c/o A T & T, Bis-P, 195 Broadway, New York City
Rossi, Dominic S., Syracuse University, Syracuse
Rusch, Reuben R., SUNY at Albany

Saam, Charles A., Chenango Valley Central School, 768 Chenango St., Binghamton
Salter, Ruth, New York State Education Department, Albany
Schmitt, John A., Boston College, Chestnut Hill, Mass.
Schumann, Mother M., Manhattanville College, Purchase
Schutte, Alfred J., Mineola High School, Garden City Park
Sefein, Naim A., SUC Fredonia
Seifert, Kenneth R., Cornell University, Ithaca
Seligson, Henry, 3344 Giles Pl., Bronx
Shafer, Dave A., Herricks Schools, New Hyde Park
Shaffen, Marcia, 154 Harvey Dr., Lancaster

Shapiro, Morris, New York State Education Department, Albany
Shands, Richard E., Guilderland Central School District, Guilderland
Shea, James G., New York University, New York City
Sipay, Edward R., SUNY at Albany
Sister M. Vivian, Rosary Hill College, 4380 Main St., Buffalo
Sister Mary Geraldine, St. Finbar School, 1825 Bath Ave., Brooklyn
Skalski, John, Fordham University, New York City
Slakter, Malcolm J., SUNY at Buffalo
Smink, Jay, Department of Public Instruction School System or Affiliation, Harrisburg, Pa.
Smith, Esther C., New York State Education Department, Albany
Smith, Livingston, New York State Education Department, Albany
Smith, Manon B., SUC Plattsburgh
Smith, Stanley V., U.S. Office of Education, Washington, D.C.
Smith, Stuart E., Alfred University, Alfred
Spain, Clarence J., Schenectady Public Schools, 108 Union St., Schenectady
Sparacio, Stephen, Baldwin Sr. High School, Baldwin
Spector, Bertram, N.Y. Institute of Technology, Wheatley Rd., Old Westbury
Sproule, Joseph R., Brighton School District, Rochester
Stambaugh, Ben F., Jr., Cornell University, Ithaca
Stauffer, Arthur L., SUC Fredonia
Steltenpohl, Elizabeth H., Manhattanville College, Purchase
Stewart, Lucille M., 254 N. Main—Apt. C4, Spring Valley
Stiglmeier, John, New York State Education Department, Albany
Stillman, Nathan, Yeshiva University, New York City
Stolper, Richard L., SUC Geneseo
Stubbins, George T., James A. Farley Jr. High School, Stony Point
Swenson, Virginia M., New York State Education Department, Albany
Tawyzer, Harold J., Hofstra University, Hempstead
Taylor, Marvin, Queens College, Flushing
Testo, Joseph J., SUNY at Albany
Thomas, George, New York State Education Department, Albany
Tonne, Herbert A., SUNY at Albany
Treffinger, Donald J., Cornell University, Ithaca
Trusty, Francis M., University of Rochester, Rochester
Tyler, Terrence, Board Cooperative Education Services, Wyoming County, Attica
Vanderhooz, William H., SUC Geneseo
Van Dyck, Richard Mrs., 11 Burnhams Pl., Delmar
Van Ryn, Mike, New York State Education Department, Albany
Varty, Jonathan W., 149 Brixton Rd., Garden City
Vieweg, William E., SUC Oneonta
Voris, George A., SUC Oneonta
Wade, Durlyn E., SUNY at Geneseo
Wadsworth, Barry J., SUNY at Albany
Walker, Charles E., University of Rochester, Rochester
Walsh, Marion V., Canandaigua Elementary School, Canandaigua
Walsh, Norman J., Columbia University, New York City
Walter, Joan M., New York State Education Department, Albany
Watson, Walter S., Cooper Union, The Cooper Square, New York City
Weber, Gerald J., U.S. Department of Health, Education and Welfare, Washington, D.C.
Wedekind, Carl E., New York State Education Department, Albany
Weir, Roger M., Brighton Central Schools, Rochester
Weser, Helen, Center on Urban Education, 16 West 16th St., New York City
Whalen, Thomas, Stoughton Public Schools, Stoughton, Mass.
White, Mary Alice, Columbia University, New York City
Wightman, Lawrence, 50 Ridge Rd., Ludlowville
Wilcox, John, SUC Oneonta
Wiley, W. Deane, New York University, New York City
Williams, Clarence M., University of Rochester, Rochester
Williamson, Clayton, P.O. Box 50, Freeport
Wilson, Lois, New York State Teachers Association, 152 Washington Ave., Albany
Winans, S. David, New Jersey State Education Department, 225 W. State St., Trenton, N.J.
Wing, Richard L., 15 Oak Rd., Montrose
Wirth, Janina W., New York State Education Department, Albany
Wohlferd, Gerald, New York State Education Department, Albany
Wolf, Richard I., SUNY at Albany
Wright, Wilber H., SUC Geneseo
Wrightstone, J. Wayne, Board of Education New York City, 110 Livingston Ave., New York City
Young, William E., New York State Education Department, Albany
Zimmerman, Gary E., SUNY at Buffalo